

APL HISTOGRAM, DENSITY ESTIMATION
AND
PROBABILITY PLOTTING ROUTINES

Dennis Roy Hutchinson

NAVAL POSTGRADUATE SCHOOL

Monterey, California



THESIS

APL HISTOGRAM, DENSITY ESTIMATION
AND
PROBABILITY PLOTTING ROUTINES

by

Dennis Roy Hutchinson

December 1976

Thesis Advisor:

P. A. W. Lewis

Approved for public release; distribution unlimited.

T176093

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS
BEFORE COMPLETING FORM

1. REPORT NUMBER		2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) APL Histogram, Density Estimation and Probability Plotting Routines			5. TYPE OF REPORT & PERIOD COVERED Master's Thesis; December 1976
7. AUTHOR(s) Dennis Roy Hutchinson			6. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940			8. CONTRACT OR GRANT NUMBER(s)
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940			10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Naval Postgraduate School Monterey, California 93940			12. REPORT DATE December 1976
			13. NUMBER OF PAGES 95
			15. SECURITY CLASS. (of this report) Unclassified
			15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.			
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)			
18. SUPPLEMENTARY NOTES			
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Data Analysis Listing; Sample Statistics (APL) (A Programming Language) Normal Plotting Histograms Exponential Plotting Jackknifing Density Estimation Sectioning Probability Plotting			
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) This paper introduces several data analysis routines that were designed for interactive use with APL (A Programming Language) and placed in the APL user library at the Naval Postgraduate School. Specifically, histograms, density estimation and probability plotting routines are both explained in detail and demonstrated with actual data. In addition, applications and limitations on each of the			

routines are explored. And, the combined routines give the general user an extensive tool to analyze either discrete or continuous data.

APL HISTOGRAM, DENSITY ESTIMATION
AND
PROBABILITY PLOTTING ROUTINES

by

Dennis Roy Hutchinson
Captain, United States Army
B.S., United States Military Academy, 1969

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the
NAVAL POSTGRADUATE SCHOOL
December 1976

ABSTRACT

This paper introduces several data analysis routines that were designed for interactive use with APL (A Programming Language) and placed in the APL user library at the Naval Postgraduate School. Specifically, histograms, density estimation and probability plotting routines are both explained in detail and demonstrated with actual data. In addition, applications and limitations on each of the routines are explored. And, the combined routines give the general user an extensive tool to analyze either discrete or continuous data.

TABLE OF CONTENTS

I.	INTRODUCTION - - - - -	8
II.	HISTOGRAM ROUTINE - - - - -	10
	A. DESCRIPTION - - - - -	10
	B. USAGE WITH TELEPHONE DATA 1 AND TELEPHONE DATA 2 OFFLINE; ALL DATA, ECDF, AND TITLE - - - - -	15
	C. USAGE WITH TELEPHONE DATA 1 AND TELEPHONE DATA 2 ON LINE; CONDITIONAL DATA BETWEEN 2 AND 140, ECDF, AND TITLE -	20
III.	LISTING ROUTINE - - - - -	25
	A. DESCRIPTION - - - - -	25
	B. USAGE WITH TELEPHONE DATA 1 AND TELEPHONE DATA 2 OFFLINE - - - - -	26
IV.	SECTIONING ROUTINE - - - - -	36
	A. DESCRIPTION - - - - -	36
	B. USAGE WITH TELEPHONE DATA 1 - - - - -	38
	C. INTERPRETATION OF RESULTS - - - - -	40
V.	JACKKNIFE ROUTINE - - - - -	41
	A. DESCRIPTION - - - - -	41
	B. USAGE WITH TELEPHONE DATA 1 - - - - -	44
	C. INTERPRETATION OF RESULTS - - - - -	46
	D. USAGE WITH COST OVERRUN DATA - - - - -	46
VI.	EXPONENTIAL PLOTTING ROUTINE - - - - -	49
	A. DESCRIPTION - - - - -	49
	B. USAGE WITH TELEPHONE DATA 1 - - - - -	49
	C. USAGE WITH RANDOM GENERATED EX- PONENTIALLY DISTRIBUTED SAMPLE WITH MEAN SAME AS TELEPHONE DATA 1 - - - -	50

VII.	NORMAL PLOTTING ROUTINE	- - - - -	53
A.	DESCRIPTION	- - - - -	53
B.	USAGE WITH COST OVERRUN DATA	- - - - -	54
C.	USAGE WITH NORMAL SAMPLE GENERATED WITH MEAN AND VARIANCE THE SAME AS COST OVERRUN DATA	- - - - -	54
D.	USAGE WITH COEFFICIENT OF VARIATION DATA OBTAINED FROM USING SECTIONING ROUTINE	- - - - -	57
VIII.	THE INDEPENDENCE AND MARKOV CHAIN HYPOTHESES FOR THE TELEPHONE DATA	- - - - -	59
IX.	DOCUMENTATION ON ROUTINES	- - - - -	73
A.	LOCATION IN APL LIBRARY	- - - - -	73
B.	WORKSPACE LOADING PROCEDURES	- - - - -	73
C.	ROUTINE LISTING	- - - - -	74
	1. Author Created Routines	- - - - -	75
	2. Adapted from Fortran Library Routine HISTG/F	- - - - -	75
	3. Borrowed Routines to Supplement Author Created Routines	- - - - -	75
X.	COMPUTER LISTING OF ALL ROUTINES	- - - - -	76
	LIST OF REFERENCES	- - - - -	94
	INITIAL DISTRIBUTION LIST	- - - - -	95

LIST OF FIGURES

Figure

1.	Histogram and descriptive statistics of telephone data 1 using routine HIST - - - - -	18
2.	Histogram and descriptive statistics of telephone data 2 using routine HIST - - - - -	19
3.	Histogram and descriptive statistics of telephone data 1 between 2 and 140 using routine HIST - - - - -	23
4.	Histogram and descriptive statistics of telephone data 2 between 2 and 140 using routine HIST - - - - -	24
5.	(a,b,c) Listing of telephone data 1 using routine HISTLIST - - - - -	28-30
6.	(a,b,c,d) Listing of telephone data 2 using routine HISTLIST - - - - -	31-34
7.	Sectioning of telephone data 1 into 16 sections with descriptive statistics using routine HISTS - - - - -	39
8.	Jackknifing of telephone data 1 into 16 groups with descriptive statistics using routine HISTJACK - - - - -	45
9.	Jackknifing of cost overrun data into 22 groups with descriptive statistics using routine HISTJACK - - - - -	48
10.	Exponential plot of telephone data 1 using routine EXPONP - - - - -	51
11.	Exponential plot of exponential sample generated with mean same as telephone data 1 using routine EXPONP - - - - -	52
12.	Normal plot of cost overrun data using routine NORMP - - - - -	55
13.	Normal plot of normal sample generated with mean and variance same as cost overrun data using routine NORMP - - - - -	56
14.	Normal plot of 16 coefficient of variation values (obtained from section routine HISTS of telephone data 1) using routine NORMP - - - - -	58

15.	Histogram and descriptive statistics of the runs of one using routine HIST	- - - - -	64
16.	Histogram and descriptive statistics of the runs of zero using routine HIST	- - - - -	65
17.	Listing of the runs of one using routine HISTLIST	- - - - -	66
18.	(a,b,c) Listing of the runs of zero using routine HISTLIST	- - - - -	67-69

ACKNOWLEDGMENTS

Particular acknowledgment is due to two individuals. Professor P.A.W. Lewis, with his extensive knowledge in the area of data analysis, was instrumental in the initial phase of this thesis. And, with his continual guidance the methodology and procedures to be used were successfully implemented in the computer routines created.

Professor Richard W. Butterworth contributed greatly with his extensive knowledge of all aspects of APL. His cooperation and willingness to assist resulted in the efficient and extensive use of all current features available for APL at the Naval Postgraduate School.

I. INTRODUCTION

The Naval Postgraduate School acquired APL (A Program-ming Language) from IBM in 1974. Since that time more and more students and faculty have become familiar with the extensive and efficient capabilities of APL and have been putting these features to good use. With the acquisition of APL came several extensive library routines that are both well documented and varied in scope. However, on close examination of these library routines it was found that statistics and data analysis were areas where some additions would be particularly useful.

Because of the efficiency and ease of APL in manipulating vectors, matrices and arrays, it is ideal for use in the area of data analysis. After a complete and thorough screening of the existing APL library routines pertaining to data analysis, it was found that by adding six additional data analysis routines to the present library, the Naval Postgraduate School could enhance its present APL capability and provide the student and general user with a more varied and flexible tool for analyzing data.

To this end the purpose of this thesis will be (1) to completely describe the six data analysis routines added to the APL library, (2) to explain the features and capabilities of each of the routines and (3) to demonstrate the use of each of the routines with "real world data".

The data to be used in this paper has come from two different sources. The first source of data was from tests performed jointly by IBM Germany and the German Public Telephone Network on errors in transmission of binary data on telephone lines (Lewis & Cox, 1966). From this source two sets of data are used and each data set contains the times between errors in binary bits transmitted over telephone lines. The first data set contains 672 elements (times-between-errors: actually number of bits between errors) and will hereby be referred to as "telephone data 1". The second data set contains 736 elements and will be referred to as "telephone data 2". The second source of data was obtained from percent overrun or underrun on selected military contracts during the year 1950 (Dixon, 1973). This data set contains 22 elements and will be referred to as "cost overrun data".

II. HISTOGRAM ROUTINE

A. DESCRIPTION

The first routine to be presented is the histogram routine which is used for estimating from given data the probability density function $f(x)$ of a continuous random variable. The current APL library has several small histogram routines that are general in nature but lack the overall detail necessary for good data analysis. For this reason HIST (histogram routine) was created. HIST represents the adaptation and modification of the fortran library version of HISTG/F, which was developed at N.P.S. by D. R. Robinson under the guidance of Professor P.A.W. Lewis. By modifying and adapting HISTG/F to APL the power and efficiency of the APL language could be put to full use.

A complete description of how HIST operates is contained in the variable HISTHOW. If the users APL workspace is properly loaded (see section IX.B. for workspace loading procedures) all that is necessary is to type HISTHOW. The user then receives the following printed response on the terminal:

HISTHOW

SYNTAX HIST

HIST ALLOWS YOU TO INTERACTIVELY OBTAIN A HISTOGRAM OF YOUR DATA ALONG WITH A SET OF BASIC DESCRIPTIVE STATISTICS. IN ADDITION, HIST HAS THE FOLLOWING CAPABILITIES WHICH ALLOW YOU:

- (1) THE OPTION OF A TITLE FOR YOUR HISTOGRAM
- (2) THE OPTION OF DISPLAYING A SMOOTHED EMPIRICAL DENSITY FUNCTION OVER THE HISTOGRAM
- (3) THE OPTION OF SCALING AND SELECTING THE NUMBER OF CELLS FOR YOUR HISTOGRAM
- (4) THE OPTION OF SELECTING AN INTERVAL AND PERFORMING A HISTOGRAM ON ALL THE DATA POINTS OR CONDITIONALLY SELECTING AN INTERVAL IN THE RANGE OF THE DATA.
- (5) THE OPTION OF HAVING YOUR OUTPUT APPEAR ON THE OFFLINE PRINTER OR ON YOUR TERMINAL

WHEN YOU TYPE HIST YOU WILL BE ASKED TO DO THE FOLLOWING:

- (1) ENTER YOUR DATA IN VECTOR FORM - YOU CAN TYPE YOUR DATA IN SINGLY OR YOU CAN TYPE THE NAME OF A VARIABLE THAT HAS YOUR DATA IN IT. YOU MUST ENSURE THAT YOU HAVE AT LEAST 10 DATA POINTS IN YOUR VECTOR AND THAT THERE IS SOME DIFFERENCES IN THE DATA POINTS (MAX SIZE OF INTEGER VECTOR IS APPROX. 2500 , MAX SIZE OF REAL VECTOR IS 2000). AFTER YOU HAVE ENTERED YOUR DATA YOU WILL BE ASKED
- (2) IF YOU DESIRE A SMOOTHED EMPIRICAL DENSITY FUNCTION OR NOT. THE EMPIRICAL DENSITY FUNCTION WHEN PLOTTED GIVES ESSENTIALLY A MORE EXACT PICTURE OF THE DATA THAN DOES THE HISTOGRAM ALONE, ALTHOUGH THIS FEATURE IS SLIGHTLY BLURRED BY THE PRECISION WHICH CAN BE OBTAINED WITH THE APL BALL (THE APL FINE PLOT IS NOT PRESENTLY AVAILABLE ON THE NPS SYSTEM). THE SMOOTHED EMPIRICAL DENSITY IS DEFINED BY THE RELATION (LEWIS,LIU,ROBINSON, AND ROSENBLATT,1975; ROSENBLATT,1956)

$$\bar{F}(Z) = \frac{1}{N} \sum_{I=1}^N \frac{W((X - Z) \div B(N))}{N \times B(N)}$$

WHERE N IS THE NUMBER OF DATA POINTS, B(N) IS A BAND-WIDTH FUNCTION,

$$B(N) = \text{RANGE} \div \text{SQRT}(N)$$

AND W IS A WEIGHT FUNCTION,

$$\begin{aligned} W(Z) &= 0 && \text{IF } |Z| > 1 \\ &= 1 - |Z| && \text{OTHERWISE} \end{aligned}$$

$\bar{F}(Z)$ IS COMPUTED FOR VALUES OF Z BETWEEN THE MAXIMUM AND THE MINIMUM OF THE SAMPLE AND PLOTTED OVER THE HISTOGRAM USING THE SYMBOL -F-. THE RELATIVE FREQUENCY MARKS ON THE LEFT OF THE OUTPUT REFER TO THE HISTOGRAM, AND NOT TO THE DENSITY FUNCTION. AFTER THIS QUERY YOU WILL BE ASKED

- (3) IF YOU DESIRE TO TITLE YOUR HISTOGRAM. IF YOU ELECT TO TITLE YOUR HISTOGRAM, SIMPLY TYPE YOUR TITLE, ENSURING THAT YOUR TITLE IS MORE THAN ONE CHARACTER IN LENGTH. IF NO TITLE IS DESIRED JUST HIT THE CARRIAGE RETURN. AFTER THE TITLE QUERY YOU WILL BE ASKED
- (4) IF YOU WANT TO SET YOUR OWN SCALE AND THE NUMBER OF CELLS. YOUR RESPONSE MUST BE A VECTOR OF 3 ELEMENTS THE FIRST ELEMENT IS THE NUMBER OF CELLS YOU DESIRE, THIS MUST BE AN INTEGER BETWEEN 10 AND 28, THE SECOND ELEMENT IS THE LEFT SCALE POINT AND THE THIRD ELEMENT IS THE RIGHT SCALE POINT (HIST DOES NOT REQUIRE THAT YOUR INTERVAL BE DIVISIBLE BY THE NUMBER OF CELLS). IF YOU WANT HIST TO AUTOMATICALLY SCALE AND PICK THE CELLS YOU SHOULD TYPE THE VECTOR 0 0 0. AFTER YOU HAVE SELECTED YOUR SCALING TECHNIQUE YOU WILL BE ASKED
- (5) IF YOU WANT DATA POINTS NOT INSIDE THE SCALE LIMITS INCLUDED IN THE HISTOGRAM ROUTINE. MOST HISTOGRAMS LUMP DATA POINTS THAT FALL OUTSIDE THE SCALE LIMITS IN THE END CELLS. HOWEVER, HIST GIVES YOU THE OPTION OF INCLUDING THEM OR EXCLUDING THEM, I.E. OF OBTAINING A HISTOGRAM FOR THE CONDITIONAL DENSITY. AFTER YOUR RESPONSE TO THIS QUERY YOU WILL BE ASKED
- (6) IF YOU WANT YOUR OUTPUT TO APPEAR ON THE OFFLINE PRINTER OR ON YOUR TERMINAL. IF YOU SELECT THE OFFLINE PRINTER THE NEXT RESPONSE YOU WILL RECEIVE ON YOUR TERMINAL IS - HISTOGRAM SENT TO PRINTER -. THIS RESPONSE WILL TAKE SEVERAL SECONDS AND AFTER IT IS RECEIVED YOUR TERMINAL IS FREE FOR FURTHER USE. HOWEVER, IF YOU ELECTED TO HAVE YOUR HISTOGRAM PRINTED ON YOUR TERMINAL THE PRINTING WOULD BEGIN IN JUST A FEW SECONDS BUT WOULD TAKE BETWEEN 5 AND 10 MINUTES TO COMPLETE.

THE FOLLOWING BASIC DESCRIPTIVE STATISTICS ARE COMPUTED AND PRINTED OUT BY HIST.

MEAN, MEDIAN, TRIMEAN, MIDMEAN, MODE
GEOMETRIC AND HARMONIC MEANS (POSITIVE SAMPLES ONLY)
VARIANCE, STANDARD DEVIATION, COEFFICIENT OF VARIATION,
RANGE AND MIDSPREAD
THIRD AND FOURTH CENTRAL MOMENTS, COEFFICIENTS OF SKEW-
NESS AND KURTOSIS
MAXIMUM, MINIMUM AND 5 SAMPLE QUANTILES

IN ADDITION, THE MEAN IS DISPLAYED ON THE HISTOGRAM BY A VERTICAL COLUMN OF -M- AND THE QUARTILES BY COLUMNS OF DOTS.

INTERPRETING THE OUTPUT

THE DEFINITIONS OF THE BASIC STATISTICS COMPUTED BY HIST ARE LISTED BELOW. PAGE NUMBER REFERENCES ARE TO THE CRC STANDARD MATH TABLES, 19TH EDITION (1971).

MEAN AVERAGE OF THE SAMPLE (P 554).

MEDIAN MID-VALUE OF THE SAMPLE, IF THERE ARE AN ODD NUMBER OF SAMPLE POINTS, OR THE AVERAGE OF THE TWO MIDDLE VALUES FOR AN EVEN NUMBER OF POINTS (P 555)

SAMPLE QUANTILES THE $Q(1)=.25$, $Q(2)=.50$, AND $Q(3)=.75$ POPULATION QUANTILES ARE THE SOLUTION TO THE EQUATION $\text{PROB}(X \leq X(Q(I))) = Q(I)$ $I=1,2,3$. THE SAMPLE QUANTILES, WHICH ESTIMATE THE POPULATION QUANTILES ARE, THE J TH ORDERED VALUE IN THE SAMPLE, WHERE $J = [Q(I) \times N] + 1$. WHERE N = SAMPLE SIZE.

TRIMEAN $0.25 \times (Q(1) + 2Q(2) + Q(3))$, WHERE THE $Q(I)$ ARE THE QUANTILES.

MIDMEAN THE AVERAGE OF ALL THE SAMPLE VALUES BETWEEN THE UPPER AND LOWER QUANTILES.

MODE THE DATA POINT THAT OCCURS MOST OFTEN (IF ALL THE DATA POINTS ARE DIFFERENT OR IF THERE ARE MORE THAN 300 DATA POINTS THE MODE WILL NOT BE PRINTED. IF TWO OR MORE MODES OCCUR HIST WILL PRINT THE FIRST MODE.)

MIDRANGE AVERAGE OF THE MAXIMUM AND MINIMUM.

GEOMETRIC (P 554).
MEAN

HARMONIC (P 555).
MEAN

VARIANCE (P 557). UNBIASED ESTIMATORS FOR VARIANCE AND STANDARD DEVIATION ARE USED.

STANDARD (P 557).
DEVIATION

COEFFICIENT OF VARIATION = STANDARD DEVIATION $\div$ |MEAN| WHEN THE MEAN IS LESS THAN $1E-30$, THE COEFFICIENT OF VARIATION IS SET TO ZERO.

MEAN (P 556). THE AVERAGE OF THE SUM OF THE ABSOLUTE DEVIATION DIFFERENCES BETWEEN THE SAMPLE VALUES AND THE MEDIAN.

RANGE MAXIMUM - MINIMUM (P 557).

MIDSPREAD $Q(3) - Q(1)$, ALSO CALLED THE INTERQUARTILE DISTANCE.

M3 THIRD CENTRAL MOMENT. UNBIASED ESTIMATOR IS USED. (P 558)

M4 FOURTH CENTRAL MOMENT. UNBIASED ESTIMATOR IS USED. (P 558)

COEFFICIENT OF SKEWNESS $M3 \div (STD DEV)^3$

COEFFICIENT OF KURTOSIS $(M4 \div (STD DEV)^4) - 3$

BETA1 BIASED ESTIMATE OF THIRD CENTRAL MOMENT. CAN BE USED IN TESTING FOR NORMALITY. (BIOMETRIKA TABLES FOR STATISTICIANS, 1966).

BETA2 BIASED ESTIMATE OF FOURTH CENTRAL MOMENT. (BIOMETRIKA TABLES FOR STATISTICIANS, 1966).

MAXIMUM LARGEST SAMPLE VALUE.

MINIMUM SMALLEST SAMPLE VALUE.

SAMPLE THE α -QUANTILE, $X(\alpha)$, IS THE SOLUTION TO THE EQ. QUANTILES PROBABILITY $(X \leq X(\alpha)) = \alpha$.

With this complete description the general user should be able to take full advantage of HIST and put to use all its options.

B. USAGE WITH TELEPHONE DATA 1 AND TELEPHONE DATA 2, OFFLINE, ALL DATA, ECDF, AND TITLE

HIST was now used on two sets of data. Both telephone data 1 and telephone data 2 were first used with the offline printer demonstrating the title option, the empirical density function option and using the conditional option with any data points outside the designated interval being lumped into the end cells. When HIST was typed the following responses to each of the queries were entered.

HIST

ENTER DATA IN VECTOR FORM

□:

TEL DAT1

IF YOU ALSO WANT A SMOOTHED EMPIRICAL DENSITY FUNCTION ENTER A 1 . IF YOU DO NOT WANT IT ENTER A 0 .

□:

1

IF YOU WANT TO TITLE YOUR HISTOGRAM TYPE YOUR TITLE. IF YOU DO NOT WANT A TITLE JUST HIT THE CARRIAGE RETURN.

TELEPHONE DATA 1

IF YOU WANT TO SET THE NUMBER OF CELLS AND THE SCALE ENTER FIRST THE NUMBER OF CELLS (AN INTEGER BETWEEN 10 AND 28) FOLLOWED BY A SPACE AND THEN YOUR LEFT SCALE POINT FOLLOWED BY A SPACE AND THEN YOUR RIGHT SCALE POINT. HOWEVER, IF YOU WANT HIST TO AUTOMATICALLY SCALE ENTER 0 0 0 .

□:

28 0 20000

GIVEN THAT YOU HAVE SET YOUR OWN SCALE, TO INCLUDE DATA POINTS THAT MIGHT BE OUTSIDE YOUR SCALE LIMITS IN THE END CELLS, TYPE 1 . IF YOU DESIGNATED AUTOSCALE ALSO, TYPE 1 . IF HOWEVER, YOU DO NOT WANT THE DATA OUTSIDE THE SCALE LIMITS INCLUDED IN THE HISTOGRAM, TYPE 0 .

□

1

IF YOU WANT YOUR OUTPUT TO APPEAR ON THE OFFLINE PRINTER, TYPE 1 . IF YOU WANT YOUR OUTPUT TO APPEAR ON YOUR TERMINAL, TYPE 0 . (NOTE IF YOU TYPED 0 BE SURE YOUR TERMINALS CARRIAGE PAGE SETTING IS ON THE MAXIMUM WIDTH)

□:

1

HISTOGRAM SENT TO PRINTER

Note that telephone data 1 was contained in the variable TELDAT1 and that the number of cells chosen was 28 with the left scale point being 0 and the right scale point being 20,000.

After the response - HISTOGRAM SENT TO PRINTER - was received. HIST was again typed under identical conditions and telephone data 2 was entered through the variable TELDAT2.

HIST

ENTER DATA IN VECTOR FORM

□:

TELDAT2

IF YOU ALSO WANT A SMOOTHED EMPIRICAL DENSITY FUNCTION ENTER A 1 . IF YOU DO NOT WANT IT ENTER A 0 .

□:

1

IF YOU WANT TO TITLE YOUR HISTOGRAM TYPE YOUR TITLE. IF YOU DO NOT WANT A TITLE JUST HIT THE CARRIAGE RETURN.

TELEPHONE DATA 2

IF YOU WANT TO SET THE NUMBER OF CELLS AND THE SCALE ENTER FIRST THE NUMBER OF CELLS (AN INTEGER BETWEEN 10 AND 28) FOLLOWED BY A SPACE AND THEN YOUR LEFT SCALE POINT FOLLOWED BY A SPACE AND THEN YOUR RIGHT SCALE POINT. HOWEVER, IF YOU WANT HIST TO AUTOMATICALLY SCALE ENTER 0 0 0 .

□:

28 0 20000

GIVEN THAT YOU HAVE SET YOUR OWN SCALE, TO INCLUDE DATA POINTS THAT MIGHT BE OUTSIDE YOUR SCALE LIMITS IN THE END CELLS, TYPE 1 . IF YOU DESIGNATED AUTOSCALE ALSO, TYPE 1 . IF HOWEVER, YOU DO NOT WANT THE DATA OUTSIDE THE SCALE LIMITS INCLUDED IN THE HISTOGRAM, TYPE 0 .

□:

1

IF YOU WANT YOUR OUTPUT TO APPEAR ON THE OFFLINE PRINTER, TYPE 1 . IF YOU WANT YOUR OUTPUT TO APPEAR ON YOUR TERMINAL, TYPE 0 . (NOTE IF YOU TYPED 0 BE SURE YOUR TERMINALS CARRIAGE PAGE SETTING IS ON THE MAXIMUM WIDTH)

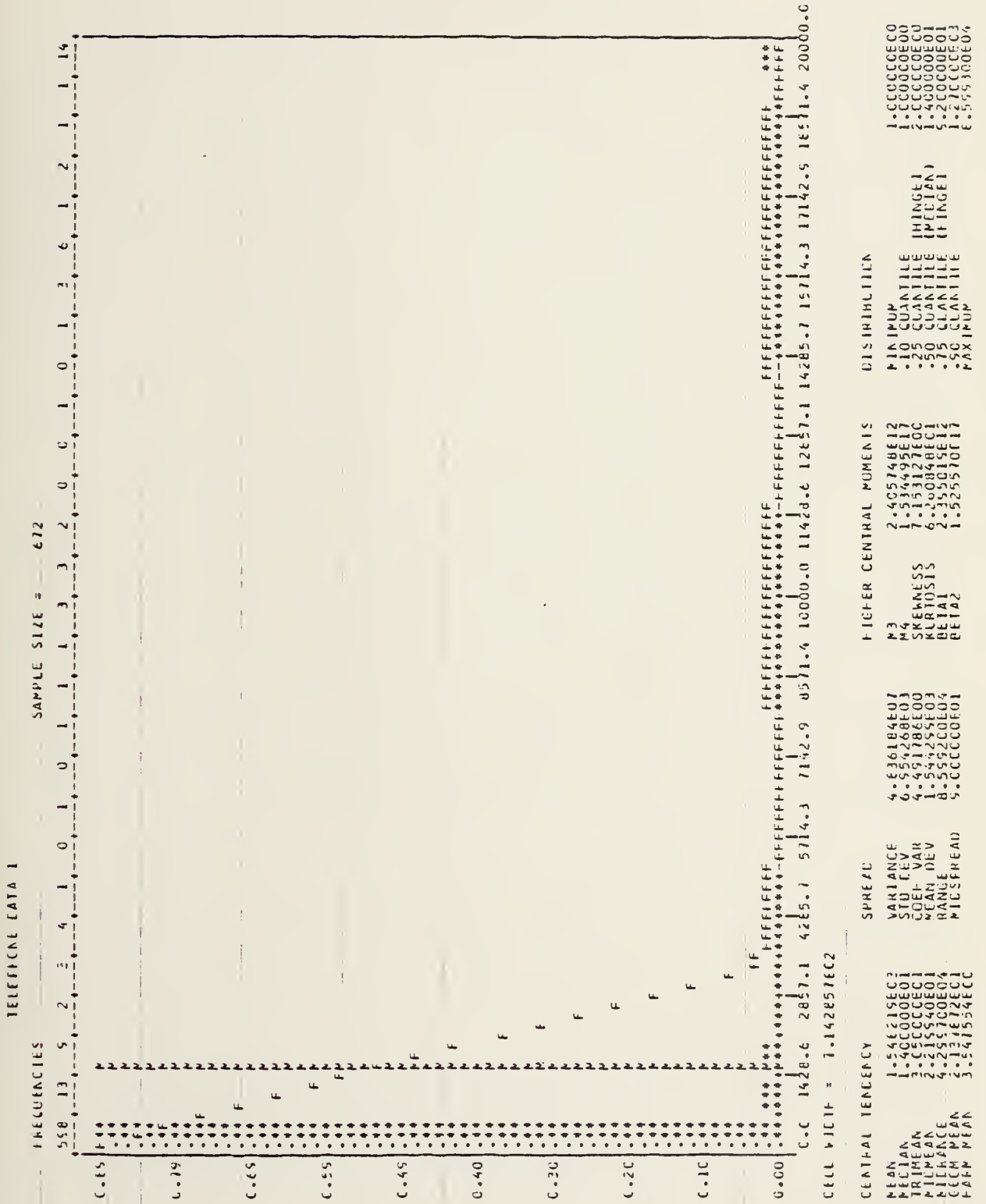
□:

1

HISTOGRAM SENT TO PRINTER

Now by looking at figure 1 (output for telephone data 1) and figure 2 (output from telephone data 2) the similarities and differences in the histograms can be compared. Without getting into specifics, the empirical density function plot seems to indicate that both sets of data are similar. However, the one time-between-errors dominate the data; a more detailed discussion of this data and its analysis is given in Section VIII.

FIGURE 1



C. USAGE WITH TELEPHONE DATA 1 AND TELEPHONE DATA 2, ON LINE; CONDITIONAL DATA BETWEEN 2 AND 140, ECDF, AND TITLE

Because both sets of data contain:

- (1) a large number of elements,
- (2) a large number of times-between-error equal to 1 (this becomes more apparent when HISTLIST is described), and
- (3) the range of the data sets is so extensive,

it would appear that the conditional option available on HIST could be used to see if the two data sets are in fact similar over a smaller interval. This in fact was done using the on line printer option, the empirical density function option, the title option and the conditional option with any data points outside the designated interval excluded from the histogram.

HIST

ENTER DATA IN VECTOR FORM

□:

TEL DAT1

IF YOU ALSO WANT A SMOOTHED EMPIRICAL DENSITY FUNCTION ENTER
A 1 . IF YOU DO NOT WANT IT ENTER A 0 .

□:

1

IF YOU WANT TO TITLE YOUR HISTOGRAM TYPE YOUR TITLE.
IF YOU DO NOT WANT A TITLE JUST HIT THE CARRIAGE RETURN.

TELEPHONE DATA 1 BETWEEN 2 AND 140

IF YOU WANT TO SET THE NUMBER OF CELLS AND THE SCALE ENTER
FIRST THE NUMBER OF CELLS (AN INTEGER BETWEEN 10 AND 28)
FOLLOWED BY A SPACE AND THEN YOUR LEFT SCALE POINT FOLLOWED
BY A SPACE AND THEN YOUR RIGHT SCALE POINT. HOWEVER, IF YOU
WANT HIST TO AUTOMATICALLY SCALE ENTER 0 0 0 .

□:

28 2 140

GIVEN THAT YOU HAVE SET YOUR OWN SCALE, TO INCLUDE DATA POINTS THAT MIGHT BE OUTSIDE YOUR SCALE LIMITS IN THE END CELLS, TYPE 1. IF YOU DESIGNATED AUTOSCALE ALSO, TYPE 1. IF HOWEVER, YOU DO NOT WANT THE DATA OUTSIDE THE SCALE LIMITS INCLUDED IN THE HISTOGRAM, TYPE 0.

□:

0

IF YOU WANT YOUR OUTPUT TO APPEAR ON THE OFFLINE PRINTER, TYPE 1. IF YOU WANT YOUR OUTPUT TO APPEAR ON YOUR TERMINAL, TYPE 0. (NOTE IF YOU TYPED 0 BE SURE YOUR TERMINALS CARRIAGE PAGE SETTING IS ON THE MAXIMUM WIDTH)

□:

0

Note that the same variable TELDAT1 is used but this time the interval was between 2 and 140. Also, the - HISTOGRAM SENT TO PRINTER - was not typed because the on-line printer (terminal) option was employed.

After the output for telephone data 1 was printed HIST was again typed and telephone data 2 was entered under identical conditions.

HIST

ENTER DATA IN VECTOR FORM

□:

TELDAT2

IF YOU ALSO WANT A SMOOTHED EMPIRICAL DENSITY FUNCTION ENTER A 1. IF YOU DO NOT WANT IT ENTER A 0.

□:

1

IF YOU WANT TO TITLE YOUR HISTOGRAM TYPE YOUR TITLE. IF YOU DO NOT WANT A TITLE JUST HIT THE CARRIAGE RETURN.

TELEPHONE DATA 2 BETWEEN 2 AND 140

IF YOU WANT TO SET THE NUMBER OF CELLS AND THE SCALE ENTER FIRST THE NUMBER OF CELLS (AN INTEGER BETWEEN 10 AND 28) FOLLOWED BY A SPACE AND THEN YOUR LEFT SCALE POINT FOLLOWED BY A SPACE AND THEN YOUR RIGHT SCALE POINT. HOWEVER, IF YOU WANT HIST TO AUTOMATICALLY SCALE ENTER 0 0 0.

□:

28 2 140

GIVEN THAT YOU HAVE SET YOUR OWN SCALE, TO INCLUDE DATA POINTS THAT MIGHT BE OUTSIDE YOUR SCALE LIMITS IN THE END CELLS, TYPE 1 . IF YOU DESIGNATED AUTOSCALE ALSO, TYPE 1 . IF HOWEVER, YOU DO NOT WANT THE DATA OUTSIDE THE SCALE LIMITS INCLUDED IN THE HISTOGRAM, TYPE 0 .

□:

0

IF YOU WANT YOUR OUTPUT TO APPEAR ON THE OFFLINE PRINTER, TYPE 1 . IF YOU WANT YOUR OUTPUT TO APPEAR ON YOUR TERMINAL, TYPE 0 . (NOTE IF YOU TYPED 0 BE SURE YOUR TERMINALS CARRIAGE PAGE SETTING IS ON THE MAXIMUM WIDTH)

□:

0

Figure 3 (output from telephone data 1 between 2 and 140) and figure 4 (output from telephone data 2 between 2 and 140) now appear quite different in shape based on the empirical density function plot. This is, again, because of the extensive range of the data (85,993 for telephone data 1 and 67,271 for telephone data 2) and the large number of times-between-error equal to one. Both sets of data are actually discrete, only occurring at multiples of 1, but as an initial analysis the data sets were treated as continuous. Thus, by employing the conditional option available on HIST differences in the two sets of data become quite apparent whereas before, the differences were not so easily detected.

FREQUENCIES

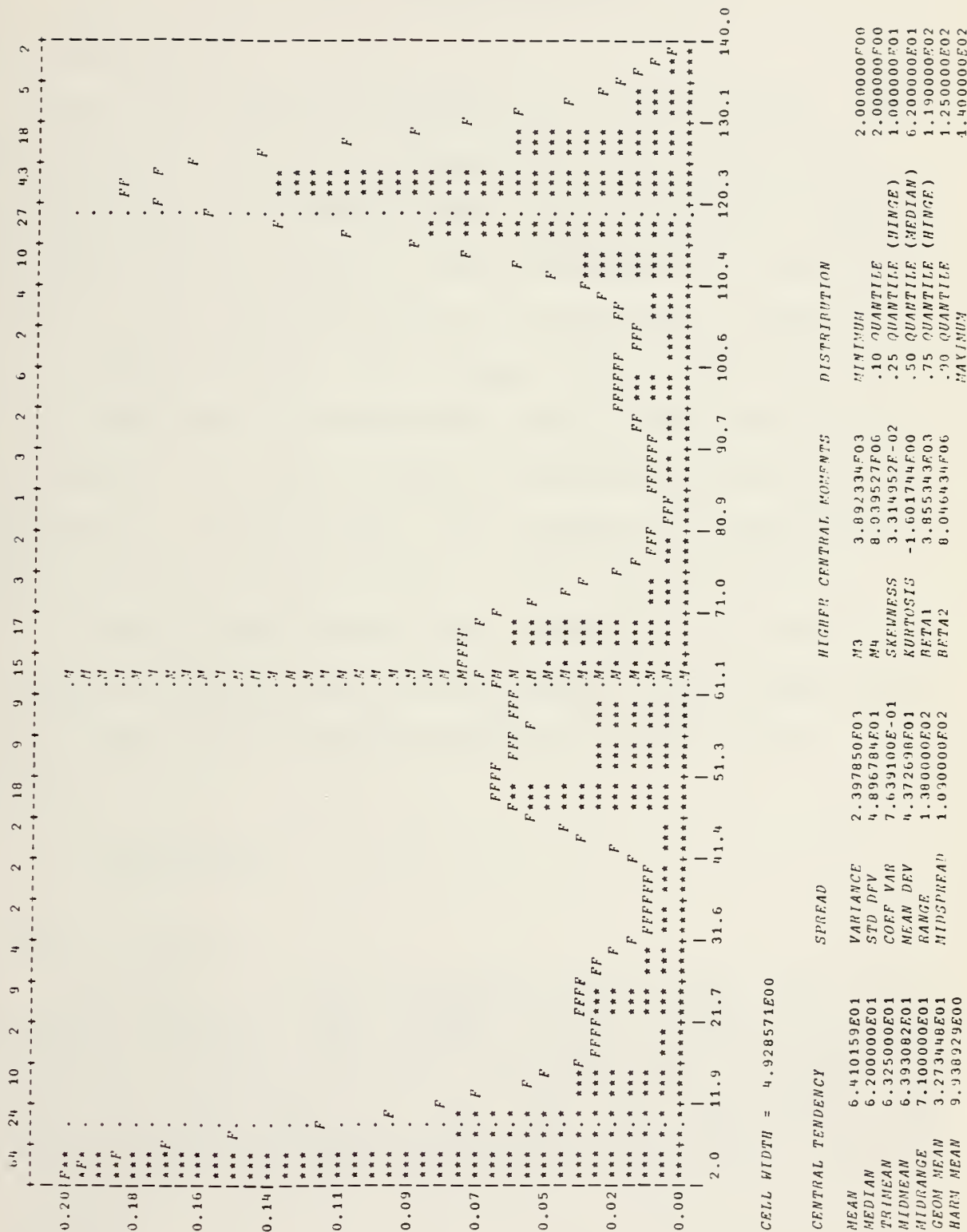
SAMPLE SIZE = 406



CELL WIDTH = 4.928571E00

CENTRAL TENDENCY		SPREAD		HIGHER CENTRAL MOMENTS		DISTRIBUTION	
MEAN	2.716995E01	VARIANCE	1.070057E03	M3	5.920433E04	MINIMUM	2.000000E00
MEDIAN	1.300000E01	STD DEV	3.271173E01	M4	5.680039E06	.10 QUANTILE	2.000000E00
TRIMEAN	1.600000E01	COEF VAR	1.203967E00	SKEWNESS	1.691386E00	.25 QUANTILE (HINGE)	4.000000E00
MIDMEAN	1.533824E01	MEAN DEV	2.165271E01	KURTOSIS	1.960633E00	.50 QUANTILE (MEDIAN)	1.300000E01
MIDRANGE	6.850000E01	RANGE	1.330000E02	BETA1	5.876758E04	.75 QUANTILE (HINGE)	3.400000E01
GEOM MEAN	1.319364E01	MIDSPREAD	3.000000E01	BETA2	5.640937E06	.90 QUANTILE	7.900000E01
HARM MEAN	6.639325E00					MAXIMUM	1.350000E02

24



III. LISTING ROUTINE

A. DESCRIPTION

The second routine presented is a listing routine. APL has a function that will automatically sort the data and print the results. However, the unique feature of HISTLIST (listing routine) is that it takes advantage of like occurrences in the data and prints the ordered data ascendingly in a compressed form. This becomes highly useful when listing a large number of data points that contain multiple occurrences. It is also a tool for finding multiplicities in supposedly continuous data, and a probability function estimating routine for data which is known to be discrete.

A complete description of how HISTLIST operates is contained in the variable HISTLISTHOW. When the user types HISTLISTHOW the following response is printed on the terminal:

HISTLISTHOW

SYNTAX HISTLIST

HISTLIST IS A HIGHLY CONVENIENT WAY TO LIST YOUR DATA. HISTLIST TAKES YOUR DATA, ORDERS IT AND COMPRESSES IT. FOR EXAMPLE, IF THREE DATA POINTS WERE ALL THE SAME VALUE HISTLIST WOULD JUST PRINT THE VALUE ONCE AND THEN PRINT THE NUMBER OF OCCURENCES OF THAT VALUE. HISTLIST WILL ALSO PRINT THE SERIAL NUMBER OF THE DATA, THE PERCENTAGE THIS SAMPLE VALUE IS TO THE WHOLE SAMPLE, AND A SMALL HISTOGRAM (STARS) SHOWING RELATIVE PERCENTAGES. EXAMPLE: 6 4 4 3 4

HISTLIST

SER. NUM.	ORDERED DATA	NUMBER OF OCCURENCES	PER CENT
1	3	1 ****	.20
2	4	3 *****	.60
5	6	1 ****	.20

HISTLIST IS IDEALLY SUITED FOR A LARGE SAMPLE THAT COULD POSSIBLY HAVE A LOT OF LIKE OCCURENCES. HISTLIST FURTHER HAS THE ADVANTAGE OF BEING USED WITH EITHER THE OFFLINE PRINTER OR THE USERS TERMINAL.

B. USAGE WITH TELEPHONE DATA 1 AND TELEPHONE DATA 2 OFFLINE

HISTLIST was used with the title option and offline printer option on both telephone data 1 and telephone data 2. When HISTLIST was typed the following responses to each of the queries were entered.

HISTLIST.

HISTLIST PRINTS THE SERIAL NUMBER OF THE COMPRESSED DATA, THE ORDERED DATA COMPRESSED, AND THE NUMBER OF LIKE OCCURENCES. ENTER YOUR DATA IN VECTOR FORM.

□:

TEL DAT1

IF YOU WANT TO TITLE YOUR DATA TYPE YOUR TITLE. IF YOU DO NOT WANT A TITLE JUST HIT THE CARRIAGE RETURN.

TELEPHONE DATA 1

IF YOU WANT YOUR OUTPUT TO APPEAR ON THE OFFLINE PRINTER TYPE 1 . IF YOU WANT YOUR OUTPUT TO APPEAR ON YOUR TERMINAL TYPE 0 .

□:

1

HISTLIST SENT TO PRINTER

After the response - HISTLIST SENT TO PRINTER - was received HISTLIST was again typed and telephone data 2 was entered.

HISTLIST
HISTLIST PRINTS THE SERIAL NUMBER OF THE COMPRESSED
DATA, THE ORDERED DATA COMPRESSED, AND THE NUMBER OF
LIKE OCCURENCES. ENTER YOUR DATA IN VECTOP FORM.
□:

TEL DAT2

IF YOU WANT TO TITLE YOUR DATA TYPE YOUR TITLE.
IF YOU DO NOT WANT A TITLE JUST HIT THE CARRIAGE
RETURN.

TELEPHONE DATA 2

IF YOU WANT YOUR OUTPUT TO APPEAR ON THE OFFLINE
PRINTER TYPE 1 . IF YOU WANT YOUR OUTPUT TO APPEAR
ON YOUR TERMINAL TYPE 0 .
□:

¹
HISTLIST SENT TO PRINTER

Looking at figure 5 (output with telephone data 1) and figure 6 (output with telephone data 2) the listings of the two data sets can be compared. It can be seen that both telephone data 1 and telephone data 2 contain a large number of multiple occurrences of the number one and the number two. In fact 19% of telephone data 1 is the number one and 24% of telephone data 2 is the number one. Also, telephone data 2 has many more multiple occurrences in the 120 to 130 range than telephone data 1. This was quickly apparent when one looked at the stars to the right of the ordered data.

FIGURE 5A

TELEPHONE DATA 1

SERIAL NUMBER	ORDERED DATA	NUMBER OF OCCURENCES	PER CENT
1	1.C00000	128	0.190
129	2.C00000	54	0.080
183	3.C00000	28	0.042
211	4.C00000	22	0.032
233	5.C00000	17	0.025
250	6.C00000	11	0.016
261	7.C00000	10	0.015
271	8.C00000	12	0.018
283	9.C00000	14	0.021
297	10.C00000	9	0.013
306	11.C00000	10	0.015
316	12.C00000	11	0.016
327	13.C00000	6	0.009
333	14.C00000	6	0.009
339	15.C00000	6	0.009
355	16.C00000	6	0.009
361	17.C00000	8	0.012
366	18.C00000	5	0.007
378	19.C00000	12	0.018
379	20.C00000	1	0.001
384	21.C00000	3	0.004
385	22.C00000	3	0.004
394	23.C00000	7	0.010
402	24.C00000	3	0.004
405	25.C00000	3	0.004
407	26.C00000	3	0.004
410	27.C00000	3	0.004
415	28.C00000	6	0.009
421	29.C00000	4	0.006
422	30.C00000	4	0.006
423	31.C00000	2	0.003
425	32.C00000	4	0.006
428	33.C00000	3	0.004
434	34.C00000	2	0.003
435	35.C00000	2	0.003
436	36.C00000	1	0.001
437	37.C00000	1	0.001
438	38.C00000	1	0.001
439	39.C00000	1	0.001
440	40.C00000	1	0.001
441	41.C00000	2	0.003
442	42.C00000	1	0.001
443	43.C00000	1	0.001
444	44.C00000	1	0.001
445	45.C00000	1	0.001
446	46.C00000	3	0.004
447	47.C00000	1	0.001
448	48.C00000	2	0.003
449	49.C00000	1	0.001
450	50.C00000	1	0.001
451	51.C00000	2	0.003
452	52.C00000	1	0.001
453	53.C00000	1	0.001
454	54.C00000	1	0.001
455	55.C00000	1	0.001
456	56.C00000	1	0.001
457	57.C00000	1	0.001
458	58.C00000	1	0.001
459	59.C00000	2	0.003
460	60.C00000	1	0.001
461	61.C00000	2	0.003
462	62.C00000	1	0.001
463	63.C00000	1	0.001
464	64.C00000	1	0.001
465	65.C00000	1	0.001
466	66.C00000	1	0.001
467	67.C00000	1	0.001
468	68.C00000	1	0.001
469	69.C00000	1	0.001
470	70.C00000	1	0.001
471	71.C00000	1	0.001
472	72.C00000	1	0.001
473	73.C00000	1	0.001
474	74.C00000	2	0.003
475	75.C00000	1	0.001
476	76.C00000	1	0.001
477	77.C00000	1	0.001
478	78.C00000	1	0.001
479	79.C00000	1	0.001
480	80.C00000	1	0.001
481	81.C00000	1	0.001
482	82.C00000	1	0.001
483	83.C00000	1	0.001
484	84.C00000	1	0.001
485	85.C00000	1	0.001
486	86.C00000	1	0.001
487	87.C00000	1	0.001
488	88.C00000	1	0.001
489	89.C00000	1	0.001
490	90.C00000	1	0.001
491	91.C00000	1	0.001
492	92.C00000	1	0.001
493	93.C00000	1	0.001
494	94.C00000	1	0.001
495	95.C00000	1	0.001
496	96.C00000	1	0.001
497	97.C00000	1	0.001
498	98.C00000	1	0.001
499	99.C00000	1	0.001
500	100.C00000	1	0.001
501	101.C00000	1	0.001
502	102.C00000	1	0.001
503	103.C00000	1	0.001
504	104.C00000	1	0.001
505	105.C00000	1	0.001
506	106.C00000	1	0.001
507	107.C00000	1	0.001
508	108.C00000	1	0.001
509	109.C00000	1	0.001
510	110.C00000	1	0.001
511	111.C00000	1	0.001
512	112.C00000	2	0.003
513	113.C00000	1	0.001
514	114.C00000	1	0.001
515	115.C00000	1	0.001
516	116.C00000	1	0.001
517	117.C00000	1	0.001
518	118.C00000	1	0.001
519	119.C00000	1	0.001

FIGURE 5B

518	1200.000000	4	0.0006
522	1221.000000	2	0.0003
524	1222.000000	2	0.0003
526	1223.000000	4	0.0006
530	1224.000000	9	0.0004
533	1228.000000	1	0.0001
534	1235.000000	1	0.0001
535	1422.000000	1	0.0001
536	1422.000000	1	0.0001
537	1422.000000	1	0.0001
538	1422.000000	1	0.0001
539	1422.000000	1	0.0001
540	1422.000000	1	0.0001
541	1422.000000	1	0.0001
542	1422.000000	1	0.0001
543	1422.000000	1	0.0001
544	1422.000000	1	0.0001
545	1422.000000	1	0.0001
546	1422.000000	1	0.0001
547	1422.000000	1	0.0001
548	1422.000000	1	0.0001
549	1422.000000	1	0.0001
550	1422.000000	1	0.0001
551	1422.000000	2	0.0001
552	1422.000000	2	0.0003
553	1422.000000	1	0.0001
554	1422.000000	1	0.0001
555	1422.000000	3	0.0004
558	1422.000000	1	0.0001
559	1422.000000	1	0.0001
560	1422.000000	1	0.0001
561	1422.000000	2	0.0001
562	1422.000000	4	0.0006
567	1422.000000	1	0.0001
568	1422.000000	1	0.0001
569	1422.000000	1	0.0001
570	1422.000000	1	0.0001
571	1422.000000	1	0.0001
572	1422.000000	1	0.0001
575	1422.000000	1	0.0001
574	1422.000000	1	0.0001
575	1422.000000	1	0.0001
576	1422.000000	1	0.0001
577	1422.000000	1	0.0001
578	1422.000000	1	0.0001
579	1422.000000	1	0.0001
580	1422.000000	1	0.0001
581	1422.000000	1	0.0001
582	1422.000000	1	0.0001
583	1422.000000	1	0.0001
584	1422.000000	1	0.0001
585	1422.000000	1	0.0001
586	1422.000000	1	0.0001
587	1422.000000	1	0.0001
588	1422.000000	1	0.0001
589	1422.000000	1	0.0001
590	1422.000000	1	0.0001
591	1422.000000	1	0.0001
592	1422.000000	1	0.0001
593	1422.000000	1	0.0001
594	1422.000000	2	0.0003
595	1422.000000	1	0.0001
597	1422.000000	1	0.0001
598	1422.000000	1	0.0001
599	1422.000000	1	0.0001
600	1422.000000	1	0.0001
601	1422.000000	1	0.0001
602	1422.000000	1	0.0001
603	1422.000000	1	0.0001
604	1422.000000	1	0.0001
605	1422.000000	1	0.0001
606	1422.000000	1	0.0001
607	1422.000000	1	0.0001
608	1422.000000	1	0.0001
609	1422.000000	1	0.0001
610	1422.000000	1	0.0001
611	1422.000000	1	0.0001
612	1422.000000	1	0.0001
613	1422.000000	1	0.0001
614	1422.000000	1	0.0001
615	1422.000000	1	0.0001
616	1422.000000	1	0.0001
617	1422.000000	1	0.0001
618	1422.000000	1	0.0001
619	1422.000000	1	0.0001
620	1422.000000	1	0.0001
621	1422.000000	1	0.0001
622	1422.000000	1	0.0001
623	1422.000000	1	0.0001
624	1422.000000	1	0.0001
625	1422.000000	1	0.0001
626	1422.000000	1	0.0001
627	1422.000000	1	0.0001
628	1422.000000	1	0.0001
629	1422.000000	1	0.0001
630	1422.000000	1	0.0001

FIGURE 5C

631	6208.000000	1	0.001
632	7614.000000	1	0.001
633	8322.000000	1	0.001
634	9015.000000	1	0.001
635	9625.000000	1	0.001
636	9866.000000	1	0.001
637	9818.000000	1	0.001
638	10154.000000	1	0.001
639	10363.000000	1	0.001
640	10451.000000	1	0.001
641	10535.000000	1	0.001
642	11260.000000	1	0.001
643	13447.000000	1	0.001
644	14365.000000	1	0.001
645	15135.000000	1	0.001
646	15264.000000	1	0.001
647	15304.000000	1	0.001
648	15347.000000	1	0.001
649	15668.000000	1	0.001
650	16260.000000	1	0.001
651	16259.000000	1	0.001
652	16361.000000	1	0.001
653	16408.000000	1	0.001
654	16317.000000	1	0.001
655	17174.000000	1	0.001
656	17667.000000	1	0.001
657	18218.000000	1	0.001
658	18649.000000	1	0.001
659	19461.000000	1	0.001
660	21848.000000	1	0.001
661	23453.000000	1	0.001
662	24692.000000	1	0.001
663	26443.000000	1	0.001
664	30574.000000	1	0.001
665	35644.000000	1	0.001
666	38003.000000	1	0.001
667	40131.000000	1	0.001
668	47120.000000	1	0.001
669	47552.000000	1	0.001
670	61710.000000	1	0.001
671	69775.000000	1	0.001
672	85993.000000	1	0.001

FIGURE 6A

TELEPHONE DATA 2

SERIAL NUMBER	ORDERED DATA	NUMBER OF OCCURENCES	PER CENT
1	1.000000	173	0.242
179	2.000000	36	0.049
215	3.000000	11	0.015
226	4.000000	6	0.008
232	5.000000	6	0.008
238	6.000000	5	0.007
243	7.000000	5	0.007
248	8.000000	4	0.005
252	9.000000	4	0.005
256	10.000000	9	0.012
259	11.000000	2	0.003
264	12.000000	3	0.004
270	13.000000	1	0.001
271	14.000000	1	0.001
272	15.000000	1	0.001
273	16.000000	4	0.005
277	17.000000	1	0.001
278	18.000000	1	0.001
279	19.000000	1	0.001
280	20.000000	3	0.004
283	21.000000	3	0.004
286	22.000000	3	0.004
288	23.000000	1	0.001
289	24.000000	3	0.004
292	25.000000	1	0.001
293	26.000000	1	0.001
294	27.000000	1	0.001
296	28.000000	1	0.001
297	29.000000	1	0.001
298	30.000000	4	0.005
302	31.000000	1	0.001
303	32.000000	4	0.005
312	33.000000	4	0.005
316	34.000000	2	0.003
319	35.000000	4	0.005
321	36.000000	4	0.005
325	37.000000	3	0.004
328	38.000000	1	0.001
329	39.000000	3	0.004
332	40.000000	3	0.004
334	41.000000	6	0.008
337	42.000000	4	0.005
338	43.000000	4	0.005
342	44.000000	3	0.004
345	45.000000	4	0.005
349	46.000000	1	0.001
350	47.000000	1	0.001
353	48.000000	3	0.004
359	49.000000	3	0.004
361	50.000000	3	0.004
366	51.000000	1	0.001
367	52.000000	1	0.001
368	53.000000	1	0.001
369	54.000000	1	0.001
370	55.000000	1	0.001
371	56.000000	1	0.001
372	57.000000	1	0.001
373	58.000000	1	0.001
374	59.000000	1	0.001
375	60.000000	1	0.001
376	61.000000	1	0.001
377	62.000000	1	0.001
378	63.000000	1	0.001
379	64.000000	3	0.004
382	65.000000	1	0.001
383	66.000000	1	0.001
384	67.000000	1	0.001
385	68.000000	1	0.001
386	69.000000	1	0.001
387	70.000000	1	0.001
388	71.000000	1	0.001
389	72.000000	1	0.001
390	73.000000	1	0.001
391	74.000000	1	0.001
395	75.000000	4	0.005
398	76.000000	1	0.001
399	77.000000	4	0.005
403	78.000000	1	0.001
404	79.000000	1	0.001
409	80.000000	3	0.004
417	81.000000	7	0.009
426	82.000000	1	0.001
433	83.000000	3	0.004
444	84.000000	3	0.004
450	85.000000	7	0.009
459	86.000000	6	0.008

FIGURE 6B

475	127.000000	3	0.0014
476	128.000000	2	0.0013
480	129.000000	4	0.0015
484	130.000000	3	0.0014
487	132.000000	3	0.0014
490	133.000000	2	0.0013
492	137.000000	1	0.0011
493	140.000000	1	0.0011
494	143.000000	1	0.0011
495	152.000000	1	0.0011
496	158.000000	1	0.0011
497	166.000000	1	0.0011
498	169.000000	2	0.0013
500	170.000000	1	0.0011
501	173.000000	2	0.0013
503	176.000000	1	0.0011
504	178.000000	1	0.0011
505	180.000000	2	0.0013
507	182.000000	1	0.0011
508	185.000000	1	0.0011
509	187.000000	1	0.0011
510	190.000000	1	0.0011
511	194.000000	1	0.0011
512	195.000000	1	0.0011
513	199.000000	1	0.0011
514	206.000000	1	0.0011
515	209.000000	1	0.0011
516	216.000000	1	0.0011
517	229.000000	1	0.0011
518	230.000000	1	0.0011
519	237.000000	1	0.0011
520	239.000000	3	0.0014
523	240.000000	2	0.0013
525	241.000000	1	0.0011
526	244.000000	2	0.0013
528	247.000000	2	0.0013
530	248.000000	3	0.0014
532	250.000000	1	0.0011
533	251.000000	1	0.0011
534	254.000000	2	0.0013
535	255.000000	2	0.0013
536	256.000000	1	0.0011
538	258.000000	1	0.0011
539	262.000000	1	0.0011
540	265.000000	1	0.0011
541	267.000000	1	0.0011
542	280.000000	1	0.0011
543	283.000000	1	0.0011
544	284.000000	1	0.0011
545	284.000000	1	0.0011
546	284.000000	1	0.0011
547	284.000000	1	0.0011
548	284.000000	1	0.0011
549	284.000000	1	0.0011
550	284.000000	2	0.0013
551	284.000000	1	0.0011
552	284.000000	1	0.0011
553	284.000000	1	0.0011
554	284.000000	1	0.0011
555	284.000000	1	0.0011
556	284.000000	1	0.0011
557	284.000000	1	0.0011
558	284.000000	1	0.0011
559	284.000000	1	0.0011
560	284.000000	1	0.0011
561	284.000000	1	0.0011
562	284.000000	1	0.0011
563	284.000000	1	0.0011
564	284.000000	2	0.0013
565	284.000000	1	0.0011
566	284.000000	1	0.0011
567	284.000000	1	0.0011
568	284.000000	1	0.0011
569	284.000000	1	0.0011
570	284.000000	1	0.0011
571	284.000000	1	0.0011
572	284.000000	1	0.0011
573	284.000000	1	0.0011
574	284.000000	1	0.0011
575	284.000000	1	0.0011
576	284.000000	1	0.0011
577	284.000000	1	0.0011
578	284.000000	1	0.0011
579	284.000000	1	0.0011
580	284.000000	1	0.0011
581	284.000000	1	0.0011
582	284.000000	1	0.0011
583	284.000000	1	0.0011
584	284.000000	1	0.0011
585	284.000000	1	0.0011
586	284.000000	1	0.0011
587	284.000000	1	0.0011
588	284.000000	1	0.0011
589	284.000000	1	0.0011
590	284.000000	1	0.0011
591	284.000000	1	0.0011
592	284.000000	1	0.0011
593	284.000000	1	0.0011
594	284.000000	1	0.0011
595	284.000000	1	0.0011
596	284.000000	1	0.0011
597	284.000000	1	0.0011
598	284.000000	1	0.0011
599	284.000000	1	0.0011
600	284.000000	1	0.0011

FIGURE 6C

595	976.000000	1	0.001
596	981.000000	1	0.001
597	1021.000000	1	0.001
598	1023.000000	1	0.001
599	1022.000000	1	0.001
600	1026.000000	2	0.003
601	1107.000000	1	0.001
602	1118.000000	1	0.001
603	1125.000000	1	0.001
604	1231.000000	1	0.001
605	1250.000000	1	0.001
606	1261.000000	2	0.003
607	1266.000000	1	0.001
608	1367.000000	1	0.001
609	1412.000000	1	0.001
610	1413.000000	1	0.001
611	1402.000000	1	0.001
612	1402.000000	1	0.001
613	1483.000000	1	0.001
614	1467.000000	1	0.001
615	1504.000000	1	0.001
616	1517.000000	1	0.001
617	1534.000000	1	0.001
618	1535.000000	2	0.003
619	1562.000000	1	0.001
620	1564.000000	1	0.001
621	1565.000000	1	0.001
622	1603.000000	1	0.001
623	1652.000000	1	0.001
624	1653.000000	1	0.001
625	1693.000000	1	0.001
626	1700.000000	1	0.001
627	1715.000000	1	0.001
628	1750.000000	1	0.001
629	1755.000000	1	0.001
630	1760.000000	1	0.001
631	1822.000000	1	0.001
632	1824.000000	1	0.001
633	1835.000000	1	0.001
634	1877.000000	1	0.001
635	1893.000000	1	0.001
636	1897.000000	1	0.001
637	1923.000000	1	0.001
638	1964.000000	1	0.001
639	1965.000000	1	0.001
640	2051.000000	1	0.001
641	2063.000000	1	0.001
642	2167.000000	1	0.001
643	2185.000000	2	0.003
644	2203.000000	1	0.001
645	2206.000000	1	0.001
646	2271.000000	1	0.001
647	2287.000000	1	0.001
648	2316.000000	1	0.001
649	2321.000000	1	0.001
650	2429.000000	1	0.001
651	2439.000000	1	0.001
652	2468.000000	1	0.001
653	2472.000000	1	0.001
654	2480.000000	1	0.001
655	2664.000000	1	0.001
656	2793.000000	1	0.001
657	2862.000000	1	0.001
658	2891.000000	1	0.001
659	2950.000000	1	0.001
660	3060.000000	1	0.001
661	3117.000000	1	0.001
662	3165.000000	1	0.001
663	3266.000000	1	0.001
664	3315.000000	1	0.001
665	3378.000000	1	0.001
666	3510.000000	1	0.001
667	3607.000000	1	0.001
668	3637.000000	1	0.001
669	3650.000000	1	0.001
670	3658.000000	1	0.001
671	3675.000000	1	0.001
672	3680.000000	1	0.001
673	3684.000000	1	0.001
674	3688.000000	1	0.001
675	3694.000000	1	0.001
676	3697.000000	1	0.001
677	3698.000000	1	0.001
678	3699.000000	1	0.001
679	3699.000000	1	0.001
680	3699.000000	1	0.001
681	3699.000000	1	0.001
682	3699.000000	1	0.001
683	3699.000000	1	0.001
684	3699.000000	1	0.001
685	3699.000000	1	0.001
686	3699.000000	1	0.001
687	3699.000000	1	0.001
688	3699.000000	1	0.001
689	3699.000000	1	0.001
690	3699.000000	1	0.001
691	3699.000000	1	0.001

FIGURE 6D

692	6385.000000	1	0.001
693	6059.000000	1	0.001
694	6816.000000	1	0.001
695	6821.000000	1	0.001
696	7307.000000	1	0.001
697	7329.000000	1	0.001
698	7146.000000	1	0.001
699	7927.000000	1	0.001
700	8039.000000	1	0.001
701	8053.000000	1	0.001
702	8253.000000	1	0.001
703	8847.000000	1	0.001
704	9200.000000	1	0.001
705	9256.000000	1	0.001
706	9517.000000	1	0.001
707	9541.000000	1	0.001
708	9562.000000	1	0.001
709	9365.000000	1	0.001
710	10020.000000	1	0.001
711	10376.000000	1	0.001
712	10637.000000	1	0.001
713	10518.000000	1	0.001
714	12042.000000	1	0.001
715	12793.000000	1	0.001
716	13012.000000	1	0.001
717	13179.000000	1	0.001
718	14092.000000	1	0.001
719	15625.000000	1	0.001
720	16877.000000	1	0.001
721	18885.000000	1	0.001
722	19273.000000	1	0.001
723	19896.000000	1	0.001
724	19849.000000	1	0.001
725	19851.000000	1	0.001
726	20840.000000	1	0.001
727	21440.000000	1	0.001
728	24373.000000	1	0.001
729	24770.000000	1	0.001
730	26278.000000	1	0.001
731	27238.000000	1	0.001
732	28133.000000	1	0.001
733	29913.000000	1	0.001
734	35667.000000	1	0.001
735	53122.000000	1	0.001
736	57271.000000	1	0.001

In addition, HISTLIST saved on printing time and paper. By printing the data in compressed form HISTLIST saved printing 448 lines (6 additional pages) in the case of telephone data 1 and 419 lines (5 additional pages) in the case of telephone data 2. Thus, HISTLIST not only gives the user more information than an ordered listing of the data, but also is cost effective in terms of printing time and paper used. Finally, note that it is not possible to look at the data in as much detail with routine HIST as with HISTLIST. If the data is continuous and there are no multiplicities, then HISTLIST gives only this information and an ordered listing of the data. The shape of the density function can best be seen (estimated) in using routine HIST.

IV. SECTIONING ROUTINE

A. DESCRIPTION

The third routine presented is the sectioning routine, HISTS. HISTS (sectioning routine) gives a way of assessing the variability of estimates of descriptive statistics from sample data. It is essential that the data be in random order.

The basic idea is as follows: Assume we have m independent observations $y_1, y_2, \dots, y_m$ of a random variable Y . The usual estimate of its mean value $\mu = E(Y)$ is the sample mean $\bar{y}$, where $\bar{y} = \sum_{i=1}^m y_i / m$. Now $\bar{y}$ is the least-squares estimate of μ , and therefore unbiased with variance $\text{var}(\bar{y}) = \sigma^2 / m$, where $\sigma^2 = \text{var}(y)$. Of course σ^2 is unknown, but we can estimate it from the data with the sample variance

$$s^2 = \frac{1}{m-1} \sum_{i=1}^m (y_i - \bar{y})^2$$

and then estimate the variance of the estimate $\bar{y}$ of μ as

$$\widetilde{\text{var}(\bar{y})} = \frac{s^2}{m} = \frac{1}{m(m-1)} \sum_{i=1}^m (y_i - \bar{y})^2$$

This is the basis for the sectioning routine: here the y_i are estimates of descriptive statistics from the m sections of the data and $\bar{y}$ is the average of the statistics

from each section. Estimates are assumed independent because the original data is assumed to be independent.

A complete description of how HISTS operates is contained in the variable HISTSHOW. When the user types HISTSHOW the following response is printed on the terminal:

HISTSHOW

SYNTAX HISTS

HISTS ALLOWS YOU TO INTERACTIVELY SECTION YOUR DATA AND ASSESS THE VARIABILITY IN EACH OF THE DESCRIPTIVE STATISTICS BY USING THE SECTIONED SAMPLE DATA.

WHEN YOU TYPE HISTS YOU WILL BE ASKED TO DESIGNATE THE NUMBER OF SECTIONS YOU DESIRE. HISTS WILL THEN TAKE THE UNORDERED DATA AND DIVIDE THE DATA INTO THE NUMBER OF SECTIONS YOU INDICATE DISCARDING ANY DATA POINTS LEFT OVER. FOR EXAMPLE, IF YOU HAVE 301 DATA POINTS AND YOU SELECT 10 SECTIONS HISTS WILL PLACE THE FIRST 30 DATA POINTS IN THE FIRST SECTION, THE SECOND 30 DATA POINTS IN THE SECOND SECTION AND SO ON UNTIL THE LAST DATA POINT IS OMITTED. YOU WILL NOW HAVE 10 SECTIONS WITH 30 DATA POINTS PER SECTION.

HISTS WOULD NOW PRINT THE FOLLOWING STATISTICS ON EACH OF THE SECTIONS: MEAN, MEDIAN, VARIANCE, STD DEV, COEF VAR, SKEWNESS, KURTOSIS, MINIMUM AND MAXIMUM. IN ADDITION, THE ABOVE STATISTICS WOULD BE PRINTED FOR THE UNSECTIONED DATA TO ALLOW FOR COMPARISONS.

FINALLY, HISTS WILL PRINT (1) THE MEAN OF THE SECTIONED DATA STATISTICS. FOR EXAMPLE, THE MEAN FOR SKEWNESS WOULD BE EACH SECTION VALUE FOR SKEWNESS SUMMED UP AND DIVIDED BY THE NUMBER OF SECTIONS. (2) THE VARIANCE AND STD DEV OF THE SECTIONED DATA STATISTICS. AND, (3) THE STD DEV DIVIDED BY THE SQUARE ROOT OF THE NUMBER OF SECTIONS, WHICH ESTIMATES THE STANDARD DEVIATION OF THE STATISTICS.

AS A RESULT, HISTS WILL GIVE YOU AN UNBIASED ESTIMATE OF THE VARIANCE OF THE SAMPLE MEAN, MEDIAN, VARIANCE, STD DEV, COEF VAR, SKEWNESS AND KURTOSIS FROM USING THE SAMPLE VARIANCE OF THE SECTIONED DATA. WITH THIS RESULT, CONFIDENCE INTERVALS CAN ALSO BE OBTAINED FOR EACH OF THE ABOVE STATISTICS, IF THE ESTIMATES FROM THE SECTIONS ARE NORMALLY DISTRIBUTED. HISTS IS BEST SUITED FOR LARGE AND MODERATE SIZED SAMPLES; FOR SMALL SAMPLES JACKKNIFING SHOULD BE CONSIDERED.

B. USAGE WITH TELEPHONE DATA 1

HISTS was now used on telephone data 1 to assess the variability in the mean, median, variance, standard deviation, coefficient of variation, skewness and kurtosis. When HISTS was typed the following responses were entered (see figure 7).

The 672 data points of telephone data 1 were broken down into 16 sections with 42 data points per section. Because of this breakdown no data points were discarded.

The unsectioned statistics printed can be compared with the values printed by HIST (figure 1) and are in fact the same. Providing that the estimates are normally distributed (this can be checked with the normal plots, described later), confidence intervals for each of the statistics (mean, median, variance, standard deviation, coefficient of variation, skewness and kurtosis) based on the t-statistic can be obtained in the following manner

$$\bar{y}_n \pm \frac{s_{\bar{y}_n}}{\sqrt{m}} t_{(1-\frac{1}{2}\alpha), (m-1)}$$

Here $\bar{y}_n$ is the mean of the sectioned data statistics (obtained from column one under summary for sectioned data); $\frac{s_{\bar{y}_n}}{\sqrt{m}}$ is the standard deviation of the sectioned data statistic divided by the square root of the number of sections (obtained from column four under summary for sectioned data); m is the number sections chosen; and, $t_{(1-\frac{1}{2}\alpha), (m-1)}$ is the $1-\frac{1}{2}\alpha$ quantile of the t-distribution with m-1 degrees of freedom.

FIGURE 7

HISTS
TYPE THE NUMBER OF SECTIONS YOU DESIRE (INTEGER
BETWEEN 2 AND 28) BE SURE TO PICK YOUR NUMBER OF
SECTIONS SO AS TO MINIMIZE THE NUMBER OF DATA
POINTS THAT WILL HAVE TO BE DISCARDED. (HISTS
PLACES THE DATA INTO THE EQUAL NUMBER OF SECTIONS
YOU INDICATE DISCARDING ANY DATA LEFT OVER)
();

16

ENTER YOUR DATA TO BE SECTIONED IN VECTOR FORM
();

TELDATE1

SECTION	MEAN	MEDIAN	VARIANCE	STD DEV	COEF VAR	SKEWNESS	KURTOSIS	MINIMUM	MAXIMUM
1	1.0526E03	8.500E08	3.4598E07	5.8820E03	5.5879E00	6.3484E08	3.7831E01	1.8000E08	3.8003E04
2	3.2135E03	1.4500E01	1.8494E08	1.3599E04	4.2328E00	5.8186E00	3.3017E01	1.0000E08	8.5993E04
3	1.7662E03	1.4500E01	4.2383E07	6.5183E03	3.6860E00	4.2886E00	1.7836E01	1.8008E00	3.5644E04
4	6.0669E02	1.1800E01	5.3412E06	2.3111E03	3.8894E00	4.3148E00	1.6488E01	1.0880E08	1.1280E04
5	1.5639E03	5.0500E01	2.1924E07	4.6824E03	2.9941E00	4.2209E00	1.8565E01	1.0000E08	2.6443E04
6	2.5343E03	5.7080E01	4.8337E07	6.3511E03	2.5861E00	3.1573E00	9.5654E00	1.0000E08	3.8974E04
7	2.6778E03	2.2800E01	7.2756E07	8.5297E03	3.1853E00	4.1587E00	1.7579E01	1.8008E08	4.7128E04
8	9.8881E02	1.8500E01	3.8282E07	6.1873E03	6.2573E00	6.4801E00	3.8995E01	1.0000E00	4.0131E04
9	1.5176E03	2.2000E01	2.0792E07	4.5599E03	3.0046E00	2.9866E00	6.8551E00	1.0000E00	1.7174E04
10	2.7682E03	1.4080E01	1.0906E08	1.8443E04	3.7726E00	4.8932E08	2.4134E01	1.0808E08	6.1718E04
11	1.9258E03	1.4080E01	5.9852E07	7.7364E03	4.0173E00	5.4134E00	2.9052E01	1.0000E00	4.7592E04
12	8.1955E02	4.9500E01	8.2895E06	2.8791E03	3.5131E00	4.5999E00	1.9765E01	1.0000E00	1.5866E04
13	2.1201E03	4.8080E00	1.2224E08	1.1056E04	5.2150E08	5.9400E00	3.3861E01	1.0000E00	6.9775E04
14	2.3062E02	1.1500E01	3.3835E05	5.7476E02	2.4923E00	3.4695E00	1.2018E01	1.8008E00	2.9620E03
15	4.3752E02	7.0000E00	5.7201E06	2.3917E03	5.4664E08	6.3983E00	3.8289E01	1.0808E08	1.5584E04
16	5.4838E02	6.5000E00	1.1340E07	3.3675E03	6.1408E00	6.4765E00	3.8964E01	1.0000E08	2.1848E04
UNSECTIONED	1.5482E03	1.4008E01	4.8362E07	6.9543E03	4.4918E00	7.1531E00	6.2608E01	1.0000E00	8.5993E04

SUMMARY FOR SECTIONED DATA

	MEAN	VARIANCE	STD DEV	STD1(SECS)*.5
MEAN	1.5482E03	8.6480E05	9.2999E02	2.3258E02
MEDIAN	2.8313E01	2.8023E02	1.6748E01	4.1850E00
VARIANCE	4.8637E07	2.6217E15	5.1283E07	1.2801E07
STD DEV	6.0664E03	1.2625E07	3.5532E03	8.8830E02
COEF VAR	4.1175E00	1.5503E00	1.2451E00	3.1128E-01
SKEWNESS	4.9343E00	1.4781E00	1.2125E00	3.0312E-01
KURTOSIS	2.4552E01	1.2484E02	1.1173E01	2.7933E00

C. INTERPRETATION OF RESULTS

As an example, a confidence interval for the coefficient of variation was obtained in the following manner. The mean value of the coefficient of variation for the 16 sections is 4.1175 (column 1). The standard deviation divided by the square root of 16 is .31128 (column 4). Using $\alpha = .05$, the t value with 15 degrees of freedom is 2.131. Thus, the 95% confidence interval for the coefficient of variation for telephone data 1 is $4.1175 \pm (.31128)(2.131)$ which is [3.454, 4.781]. Confidence intervals on the six other statistics could be obtained in the same fashion.

Again note that the use of the variance estimate from the sectioned data to give confidence intervals is based on the assumption that the estimates from the sections are independent and normally distributed. The normality will depend on the number of observations in each section, which should be kept large to induce normality. This requirement conflicts with the need to make the number of sections large to reduce the variability in the estimate of the variance of the statistics.

Another problem is that if the number of observations in each section is small, the estimates may be severely biased. This effect can be seen in figure 7: note that all of the 16 estimates of skewness from the sections are smaller than the estimate 7.1531 from the unsectioned data.

V. JACKKNIFE ROUTINE

A. DESCRIPTION

The fourth routine presented is the jackknife routine. HISTJACK (jackknife routine) is another way of assessing the variability in the estimates from sample data, and also of reducing bias in estimates of the descriptive statistics.

The jackknife procedure, like the previous sectioning method, is based on the assumption that an independent and identically distributed random sample $x_1, x_2, \dots, x_n$ have come from a population with an unknown distribution function $F_X(x)$. If we divide the sample into r groups, with each group containing the same number of elements, we can obtain estimates $\tilde{\theta}$ of the descriptive statistics, which we denote generically as θ , in the same manner as previously done with the sectioning method. The difference here is that the descriptive statistics are computed with the j^{th} group deleted $j=1,2,\dots,r$. We then let $\tilde{\theta}_{(j)}$ be the result or the descriptive statistic estimate computed with the j^{th} subgroup omitted, and $\tilde{\theta}_{a11}$ is the corresponding result or descriptive statistic estimated from the entire sample (no group omitted). The jackknife pseudo-values are then computed in the following way:

$$\tilde{\theta}_{*j} = (r)(\tilde{\theta}_{a11}) - (r-1)(\tilde{\theta}_{(j)}) \quad j = 1,2,\dots,r$$

Then we define the jackknifed estimator to be:

$$\tilde{\theta}_{*} = \frac{1}{r} \sum_{j=1}^r \tilde{\theta}_{*j}$$

The pseudo-values can be used to obtain variance estimates for $\tilde{\theta}_{*}$, and to set approximate confidence limits, using Student's t . The idea is that the pseudo-values will be approximately independent and possibly normally distributed. The jackknifed estimator $\tilde{\theta}_{*}$ is a sample average so we form an estimate s_{*}^2 of its variance given by the following relationship (Miller, 1974):

$$s^2 = \frac{\sum \tilde{\theta}_{*j}^2 - \frac{1}{r} (\sum \tilde{\theta}_{*j})^2}{r-1}$$

$$s_{*}^2 = s^2/r$$

This procedure is particularly useful if the number n of data points is small, but it must be used with care. Note, that the estimator $\tilde{\theta}_{*}$ is designed to eliminate a $1/n$ bias term in the estimator $\tilde{\theta}$.

A complete description of how HISTJACK operates is contained in the variable HISTJACKHOW. When the user types HISTJACKHOW the following response is printed on the terminal.

HISTJACKHOW

SYNTAX HISTJACK

HISTJACK ALLOWS YOU TO INTERACTIVELY JACKKNIFE YOUR DATA AND ASSESS THE VARIABILITY IN EACH OF THE STATISTICAL ESTIMATES BY USING THE SAMPLE DATA.

WHEN YOU TYPE HISTJACK YOU WILL BE ASKED TO DESIGNATE THE NUMBER OF GROUPS YOU DESIRE. HISTJACK WILL TAKE THE UNORDERED DATA AND DIVIDE THE DATA INTO THE NUMBER OF GROUPS YOU INDICATE DISCARDING ANY DATA POINTS LEFT OVER. FOR EXAMPLE, IF YOU HAVE 22 DATA POINTS AND YOU SELECT 7 GROUPS HISTJACK WILL PLACE THE FIRST 3 DATA POINTS IN GROUP 1, THE SECOND 3 DATA POINTS IN GROUP 2, AND SO ON UNTIL THE LAST DATA POINT IS OMITTED. YOU WOULD NOW HAVE 7 GROUPS WITH 3 DATA POINTS PER GROUP. IF YOU HAD ELECTED TO DO A COMPLETE JACKKNIFE, THAT IS TYPED 22, YOU WOULD NOW HAVE 22 GROUPS WITH 1 DATA POINT OMITTED PER GROUP.

HISTJACK WOULD NOW PERFORM STATISTICAL COMPUTATIONS USING THE JACKKNIFE PROCEDURE. THAT IS, BY OMITTING ONE GROUP AT A TIME, STARTING WITH THE FIRST GROUP, HISTJACK WOULD PRINT THE FOLLOWING STATISTICS: MEAN, MEDIAN, VARIANCE, STD DEV, COEF VAR, SKEWNESS, KURTOSIS, MINIMUM AND MAXIMUM. IN ADDITION, THE ABOVE STATISTICS WOULD BE PRINTED FOR THE UNGROUPED DATA TO ALLOW FOR COMPARISONS. (NOTE, THE COLUMNS GIVE THE STATISTIC ESTIMATED FROM ALL THE DATA WITH ONE GROUP MISSING, AND NOT THE PSEUDO-VALUES)

FINALLY, HISTJACK WILL PRINT (1) THE JACKKNIFE ESTIMATE (2) THE SAMPLE VARIANCE OF THE PSEUDO-VALUES DERIVED IN THE JACKKNIFE ESTIMATE (3) AND, THE ESTIMATED STD DEV OF THE JACKKNIFE ESTIMATE DIVIDED BY THE SQUARE ROOT OF THE NUMBER OF GROUPS.

AS A RESULT, HISTJACK WILL GIVE YOU AN ESTIMATE OF THE VARIANCE OF THE SAMPLE MEAN, MEDIAN, VARIANCE, STD DEV, COEF VAR, SKEWNESS AND KURTOSIS USING THE SAMPLE VARIANCE OF THE JACKKNIFED DATA. WITH THIS RESULT, CONFIDENCE INTERVALS CAN BE OBTAINED FOR EACH OF THE ABOVE STATISTICS, AGAIN ASSUMING THAT THE PSEUDO-VALUES ARE APPROXIMATELY INDEPENDENT AND NORMALLY DISTRIBUTED. HISTJACK IS BEST SUITED FOR SMALL SAMPLES.

B. USAGE WITH TELEPHONE DATA 1

HISTJACK was now used on telephone data 1 to assess the variability in the mean, median, variance, standard deviation, coefficient of variation, skewness and kurtosis. When HISTJACK was typed the following responses were entered. (see figure 8)

The 672 data points were broken down into 16 groups with 42 data points per group. Again, because of this breakdown no data points were discarded.

The ungrouped statistics printed are again the same values that were printed by HIST (figure 1). Using the jackknife method, confidence intervals for each of the statistics (mean, median, variance, standard deviation, coefficient of variation, skewness and kurtosis) can be obtained in the following manner;

$$\tilde{\theta}_* \pm (s_*) t_{(1-\frac{1}{2}\alpha), (r-1)} .$$

Here $\tilde{\theta}_*$ is the jackknife estimate of the sample data (obtained from column one under summary for jackknifed data); s_* is the jackknife estimate of the standard deviation divided by the square root of the number of groups (obtained from column four under summary for jackknifed data); r is the number of groups chosen; and, $t_{(1-\frac{1}{2}\alpha), (r-1)}$ is the $1-\frac{1}{2}\alpha$ quantile of the t-distribution with $r-1$ degrees of freedom. The basis for these assertions about the confidence intervals using the jackknifing technique is asymptotic and great care must be taken in using them.

FIGURE 8

HISTJACK
TYPE THE NUMBER OF GROUPS YOU DESIRE (INTEGER
BETWEEN 2 AND 50) BE SURE TO PICK YOUR NUMBER
OF GROUPS SO AS TO MINIMIZE THE NUMBER OF DATA
POINTS THAT WILL HAVE TO BE DISCARDED. (HISTJACK
PLACES THE DATA INTO THE EQUAL NUMBER OF GROUPS
YOU INDICATE DISCARDING ANY DATA LEFT OVER)

16

ENTER YOUR DATA TO BE JACKKNIFED IN VECTOR FORM
():
TELDATE1

GROUP	MEAN	MEDIAN	VARIANCE	STD DEV	COEF VAR	SKEWNESS	KURTOSIS	MINIMUM	MAXIMUM
1	1.5813E03	1.5000E01	4.9318E07	7.0227E03	4.4412E00	7.1746E00	6.3025E01	1.0000E00	8.5993E04
2	1.4372E03	1.4000E01	3.9338E07	6.2720E03	4.3641E00	6.5220E00	5.0690E01	1.0000E00	6.9775E04
3	1.5337E03	1.4000E01	4.0825E07	6.9875E03	4.5560E00	7.3093E00	6.4762E01	1.0000E00	8.5993E04
4	1.6110E03	1.4000E01	5.1180E07	7.1540E03	4.4408E00	6.9781E00	5.9257E01	1.0000E00	8.5993E04
5	1.5472E03	1.3500E01	5.0162E07	7.0825E03	4.5777E00	7.1494E00	6.1827E01	1.0000E00	8.5993E04
6	1.4825E03	1.3000E01	4.0893E07	6.9923E03	4.7166E00	7.3663E00	6.5160E01	1.0000E00	8.5993E04
7	1.4729E03	1.3000E01	4.6758E07	6.8380E03	4.6425E00	7.5081E00	6.8831E01	1.0000E00	8.5993E04
8	1.5855E03	1.4000E01	4.9073E07	7.0052E03	4.4183E00	7.1850E00	6.3371E01	1.0000E00	8.5993E04
9	1.5503E03	1.3500E01	5.0236E07	7.0877E03	4.5720E00	7.1592E00	6.1773E01	1.0000E00	8.5993E04
10	1.4669E03	1.4000E01	4.4376E07	6.6615E03	4.5413E00	7.4572E00	6.9618E01	1.0000E00	8.5993E04
11	1.5230E03	1.4000E01	4.7680E07	6.9050E03	4.5337E00	7.3200E00	6.5949E01	1.0000E00	8.5993E04
12	1.5968E03	1.3000E01	5.1013E07	7.1423E03	4.4729E00	7.0092E00	5.9701E01	1.0000E00	8.5993E04
13	1.5101E03	1.6000E01	4.3600E07	6.6030E03	4.3726E00	7.1493E00	6.5032E01	1.0000E00	8.5993E04
14	1.6361E03	1.4000E01	5.1446E07	7.1726E03	4.3841E00	6.9200E00	5.8526E01	1.0000E00	8.5993E04
15	1.6223E03	1.5000E01	5.1130E07	7.1506E03	4.4078E00	6.9784E00	5.9319E01	1.0000E00	8.5993E04
16	1.6149E03	1.5000E01	5.0781E07	7.1261E03	4.4128E00	7.0291E00	6.0129E01	1.0000E00	8.5993E04
UNGROUPED	1.5482E03	1.4000E01	4.8362E07	6.9543E03	4.4918E00	7.1531E00	6.2608E01	1.0000E00	8.5993E04

SUMMARY FOR JACKKNIFED DATA

(VAR*GROUPS)*.5
JACKKNIFE ESTIMATE OF STD DEV
OF MEAN OF PSEUDO-VALUES

MEAN	VARIANCE
1.5482E03	8.6480E05
1.3063E01	1.5656E02
4.8344E07	2.5453E15
7.0154E03	1.3879E07
4.5053E00	2.4262E00
7.3732E00	1.2963E01
6.7077E01	4.6806E03
2.3250E02	
3.1281E00	
1.2613E07	
9.3135E02	
3.8940E-01	
9.0012E-01	
1.7104E01	

C. INTERPRETATION OF RESULTS

To compare the confidence interval obtained for the coefficient of variation using the sectioning routine with that obtained using the jackknife routine the following was done. The jackknife estimate of the coefficient of variation for the 16 groups is 4.5053 (column 1). The jackknife estimate of the standard deviation divided by the square root of 16 is .3894 . Using $\alpha = .05$, the t value with 15 degrees of freedom is 2.131. Thus, the 95% confidence interval for the coefficient of variation for telephone data 1 is $4.5053 \pm (.3894)(2.131)$ which is [3.676, 5.335]. This compares with the confidence interval of [3.454, 4.781] using the sectioning routine described in section IV. Likewise, confidence intervals on the remaining six statistics could be obtained in a similar manner. Note that the values obtained for the skewness coefficient from the sections are now not evidently biased; of the 16 values, 7 have values below the value 7.1531 for all the data.

D. USAGE WITH COST OVERRUN DATA

To demonstrate how the complete jackknife could be used and why it is better to use when possible, the following was done. The 22 data points of the cost overrun data were used with the jackknife routine (HISTJACK). When HISTJACK was typed the data was entered in the variable YROVR and 22 was typed as the number of groups. By typing 22, which is the same as the number of data points, a complete jackknife was done.

Looking at the output from the complete jackknife (figure 9), the cost overrun data can be studied. One can note that by using the complete jackknife the mean, median, and variance of the jackknife estimate (column one under summary for jackknifed data) are the same value as the ungrouped mean, median and variance. But, also note that the coefficient of variation is less than zero which can happen when using the jackknife technique.

FIGURE 9

HISTJACK
TYPE THE NUMBER OF GROUPS YOU DESIRE (INTEGER
BETWEEN 2 AND 50) BE SURE TO PICK YOUR NUMBER
OF GROUPS SO AS TO MINIMIZE THE NUMBER OF DATA
POINTS THAT WILL HAVE TO BE DISCARDED. (HISTJACK
PLACES THE DATA INTO THE EQUAL NUMBER OF GROUPS
YOU INDICATE DISCARDING ANY DATA LEFT OVER)
[]:

22

ENTER YOUR DATA TO BE JACKKNIFE IN VECTOR FORM
[]:

YR0VR

GROUP	MEAN	MEDIAN	VARIANCE	STD DEV	COEF VAR	SKEWNESS	KURTOSIS	MINIMUM	MAXIMUM
1	1.0524E00	-1.4000F00	1.0228F02	1.0113F01	9.6101E00	7.7349E-01	7.3991E-02	-1.3600F01	2.5300F01
2	1.2048E00	-1.2000F00	1.0189F02	1.0094F01	8.3784E00	7.2904E-01	5.1529E-02	-1.3600F01	2.5300F01
3	1.3190E00	-1.2000F00	1.0089F02	1.0044F01	7.6149E00	7.0897E-01	7.9345E-02	-1.3600F01	2.5300F01
4	1.7714E00	-1.2000E00	9.1014E01	9.5401E00	5.3056E00	8.7130E-01	3.2139E-01	-1.3000F01	2.5300F01
5	9.9048E-01	-1.6000F00	1.0213F02	1.0106F01	1.0203E01	7.9519E-01	1.0417E-01	-1.3600F01	2.5300F01
6	7.6190E-01	-1.6000F00	1.0006F02	1.0003E01	1.3129E01	8.8022E-01	3.1867E-01	-1.3600F01	2.5300F01
7	1.2286E00	-1.2000F00	1.0173F02	1.0086E01	8.2096E00	7.2372E-01	5.4304E-02	-1.3600F01	2.5300F01
8	1.4381E00	-1.2000F00	9.9207F01	9.9603E00	6.9260E00	7.0632E-01	1.4054E-01	-1.3600F01	2.5300F01
9	1.5286E00	-1.2000F00	9.7491F01	9.8738F00	6.4595E00	7.2120E-01	2.0046E-01	-1.3600F01	2.5300F01
10	1.2000E00	-1.2000E00	1.0192E02	1.0095E01	8.4128E00	7.3016E-01	5.1155E-02	-1.3600E01	2.5300F01
11	1.0190E00	-1.6000E00	1.0222F02	1.0111E01	9.9216E00	7.8499E-01	8.8707E-02	-1.3600F01	2.5300F01
12	1.7429E00	-1.2000E00	9.1918F01	9.5874E00	5.5009E00	8.4272F-01	3.1785E-01	-1.3600F01	2.5300F01
13	1.6000E00	-1.2000F00	9.5863E01	9.7913E00	6.1195E00	7.4622E-01	2.4899E-01	-1.3600F01	2.5300F01
14	1.2857E00	-1.2000E00	1.0124F02	1.0062E01	7.8260E00	7.1331E-01	6.7672E-02	-1.3600E01	2.5300F01
15	6.0476E-01	-1.6000F00	9.7232F01	9.8607E00	1.6305F01	9.3010E-01	5.4286E-01	-1.3600F01	2.5300F01
16	1.9048E-01	-1.6000F00	8.4311F01	9.1821E00	4.8206F01	8.9026E-01	9.9254E-01	-1.3600E01	2.5300E01
17	6.8571E-01	-1.6000F00	9.8831E01	9.9414E00	1.4498E01	9.0629E-01	4.2160E-01	-1.3600E01	2.5300F01
18	-8.0952E-02	-1.6000F00	7.1540F01	8.4585E00	1.0449F02	4.6734E-01	-2.0707E-01	-1.3600F01	1.9600F01
19	1.1810E00	-1.6000E00	1.0202F02	1.0101F01	8.5529E00	7.3487E-01	5.0328E-02	-1.3600F01	2.5300F01
20	9.6190E-01	-1.6000E00	1.0201F02	1.0100F01	1.0500E01	8.0563E-01	1.2225E-01	-1.3600E01	2.5300F01
21	1.3333E00	-1.2000F00	1.0072F02	1.0036E01	7.5270E00	7.0754E-01	8.5165E-02	-1.3600E01	2.5300F01
22	5.8095E-01	-1.6000F00	9.6705F01	9.8339E00	1.6927E01	9.3606E-01	5.8006E-01	-1.3600F01	2.5300F01
UNGROUPED	1.0727E00	-1.4000F00	9.7420F01	9.8702F00	9.2010F00	7.8191E-01	2.2400E-01	-1.3600F01	2.5300F01

SUMMARY FOR JACKKNIFE DATA

(VARIGROUPS)*.5
JACKKNIFE ESTIMATE OF STD DEV
OF MEAN OF PSEUDO-VALUES

JACKKNIFE ESTIMATE VARIANCE

MEAN	1.0727E00	9.7420F01	2.1043E00
MEDIAN	-1.4000E00	1.8460E01	9.1652E-01
VARIANCE	9.7420F01	2.3840E04	3.2919F01
STD DEV	1.0026E01	6.7477E01	1.7513F00
COEF VAR	-1.2279E02	2.0893E05	9.7452E01
SKEWNESS	8.7458E-01	4.8496F00	4.6951E-01
KURTOSIS	4.3524E-01	2.7843F01	1.1250E00

VI. EXPONENTIAL PLOTTING ROUTINE

A. DESCRIPTION

The fifth routine presented is an exponential plotting routine. Routine EXPONP is a way of plotting the data to see if it "fits" an exponential distribution, and also to give some indication of what alternative distributions could be used if the exponential hypothesis is rejected.

A complete description of how EXPONP operates is contained in the variable EXPONPHOW. When the user types EXPONPHOW the following response is printed on the terminal.

EXPONPHOW

SYNTAX EXPONP

EXPONP ORDERS THE DATA X(I) AND COMPUTES THE EMPIRICAL LOG SURVIVER FUNCTION FOR THE DATA. THAT IS,

$$\begin{array}{c} \backslash / \\ X \\ / \backslash (I) \end{array} \quad \text{VS} \quad \begin{array}{c} | \quad | \backslash | \\ | \quad | \backslash | \\ | \_ \_ | \quad | \end{array} \quad \begin{array}{c} / \\ | \quad 1 - \\ \backslash \end{array} \quad \begin{array}{c} I \\ \text{-----} \\ N+1 \end{array} \quad \begin{array}{c} \backslash \\ | \\ / \end{array}$$

THE ORDERED DATA IS PLOTTED AGAINST THE LOG SURVIVER FUNCTION TO SEE IF THERE IS A LINEAR FIT. EXPONP ALSO ALLOWS YOU TO TITLE YOUR PLOT.

B. USAGE WITH TELEPHONE DATA 1

EXPONP was used with telephone data 1 to see if the data plotted as a relative straight line. When EXPONP was typed the following responses were entered.

EXPONP
EXPONP ORDERS THE DATA YOU GIVE AND COMPUTES THE
EMPIRICAL LOG SURVIVER FUNCTION FOR THE DATA.
A PLOT OF THE LOG SURVIVER FUNCTION FOR THE DATA
IS THEN PRINTED TO SEE IF THERE IS A LINEAR FIT.

IF YOU WANT TO TITLE YOUR PLOT TYPE YOUR TITLE.
IF YOU DO NOT WANT A TITLE JUST HIT THE CARRIAGE
RETURN.

TELEPHONE DATA 1

ENTER YOUR DATA IN VECTOR FORM

□:

TEL DAT1

Looking at figure 10 (plot of telephone data 1 using EXPONP), it was found that the data did not plot linearly from the origin, but that the data did appear somewhat linear in the tail (5,000 to 90,000 range).

C. USAGE WITH RANDOM GENERATED EXPONENTIALLY DISTRIBUTED SAMPLE WITH MEAN SAME AS TELEPHONE DATA 1

As a comparison, EXPONP was used with an exponentially generated random sample with the same mean as telephone data 1 (figure 11). As expected, this plot is, within limits of sample fluctuations, linear from the origin and in fact, what telephone data 1 would have looked like if the data was truly exponential. The quantization because of the coarseness of the APL type-ball is evident in this plot. The sample size is 672 , but not all these points can be plotted separately.

FIGURE 10

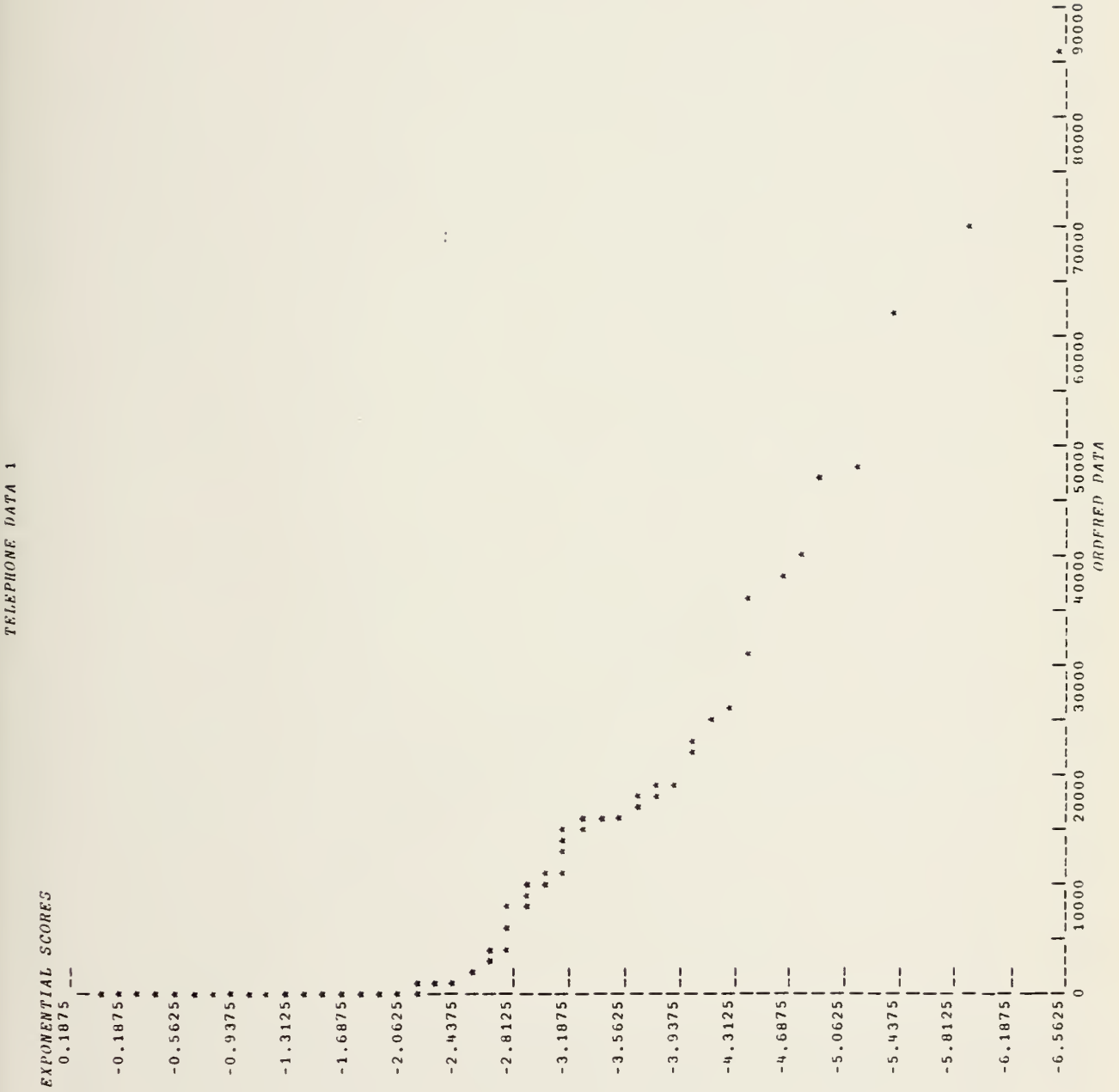
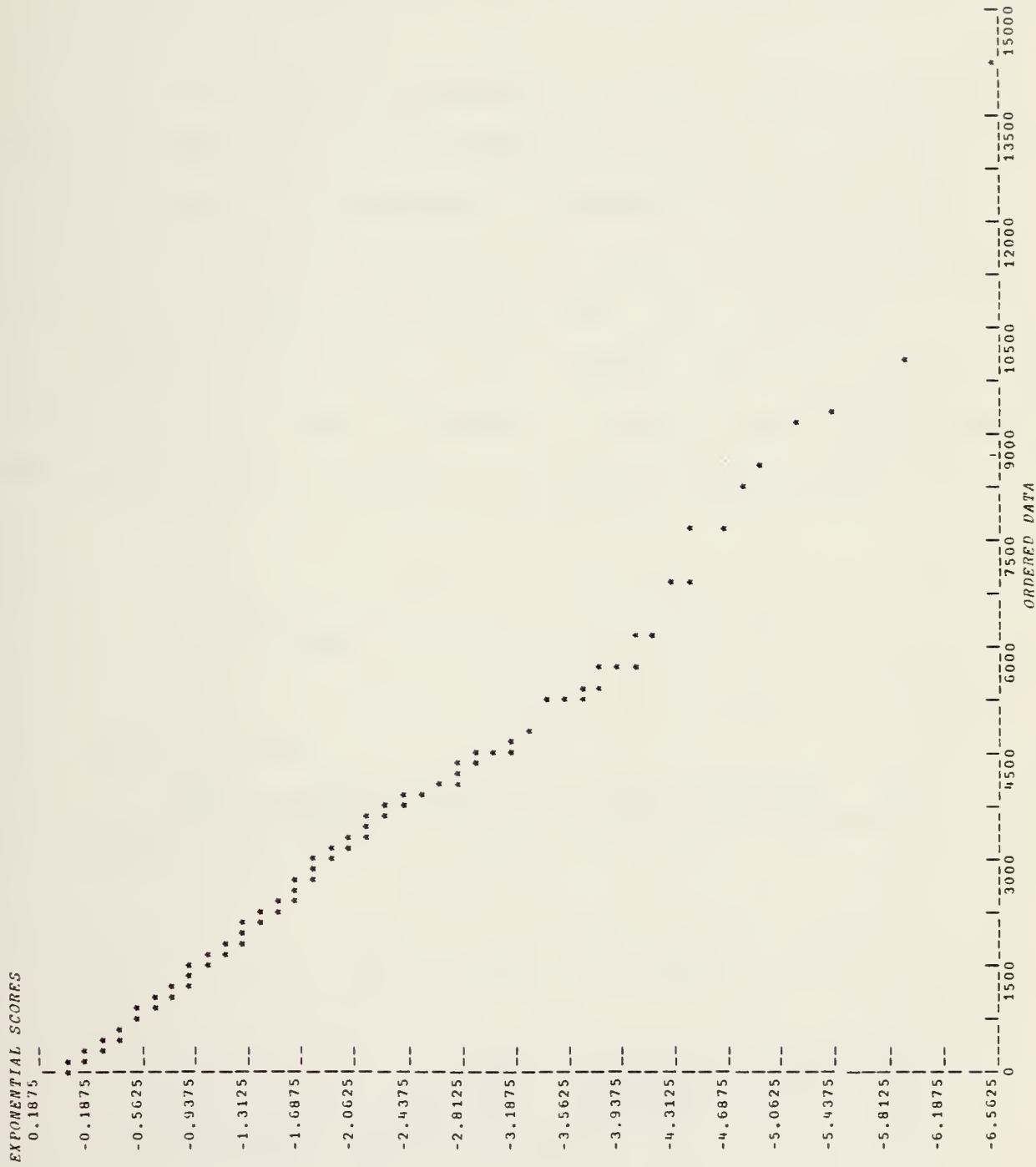


FIGURE 11

EXPONENTIALLY GENERATED RANDOM SAMPLE WITH MEAN THE SAME AS TELEPHONE DATA 1



VII. NORMAL PLOTTING ROUTINE

A. DESCRIPTION

The final routine presented is a normal plotting routine. Routine NORMP is a way of plotting the data to see if it "fits" a normal distribution. In particular one might want to look at estimates of descriptive statistics obtained from sections and groups in routines HISTS and HISTJACK.

A complete description of how NORMP operates is contained in the variable NORMPHOW. When the user types NORMPHOW the following response is printed on the terminal.

NORMPHOW

SYNTAX NORMP

NORMP ORDERS THE DATA $X(I)$ AND COMPUTES THE INVERSE OF THE UNIT NORMAL CUMULATIVE DISTRIBUTION. THAT IS,

$$\begin{array}{c} \backslash / \\ \times \\ / \backslash \end{array} (I) \quad \text{VS} \quad \begin{array}{c} T-1 / I \backslash \\ \Phi \quad | \text{-----} | \\ 1 \quad \backslash N+1 / \end{array}$$

THE ORDERED DATA IS PLOTTED AGAINST THE INVERSE OF THE UNIT NORMAL CUMULATIVE DISTRIBUTION TO SEE IF THERE IS A LINEAR FIT. NORMP ALSO ALLOWS YOU TO CONVIENTLY TITLE YOUR PLOT.

B. USAGE WITH COST OVERRUN DATA

NORMP was used with the cost overrun data to see if the data plotted as a relative straight line. When NORMP was typed the following responses were entered.

```
NORMP
NORMP ORDERS THE DATA YOU GIVE AND COMPUTES THE
INVERSE OF THE UNIT NORMAL CUMULATIVE DISTRIBUTION
FOR THE DATA. A PLOT OF THE INVERSE OF THE
UNIT NORMAL CUMULATIVE DISTRIBUTION VS THE ORDERED
DATA IS THEN PRINTED TO SEE IF THERE IS A
LINEAR FIT.
```

```
IF YOU WANT TO TITLE YOUR PLOT TYPE YOUR TITLE.
IF YOU DO NOT WANT A TITLE JUST HIT THE CARRIAGE
RETURN.
```

```
COST OVERRUNS
```

```
ENTER YOUR DATA IN VECTOR FORM
□:
```

```
YROVR
```

Note that the cost overrun data was contained in the variable YROVR . Looking at figure 12 (plot of cost overrun data using NORMP), it was found that the data did in fact plot fairly linear through the range -14 to 26 (formal tests are available; see Wilk & Gnanadesikan, 1968).

C. USAGE WITH NORMAL SAMPLE GENERATED WITH MEAN AND VARIANCE THE SAME AS COST OVERRUN DATA

As a comparison, NORMP was used with a normal sample with the same mean and variance as the cost overrun data (figure 13). As expected, this plot is very linear. But again, this plot is not that much different from that of figure 12, which gives credence to the fact that the cost overrun data might in fact be normally distributed.

FIGURE 12

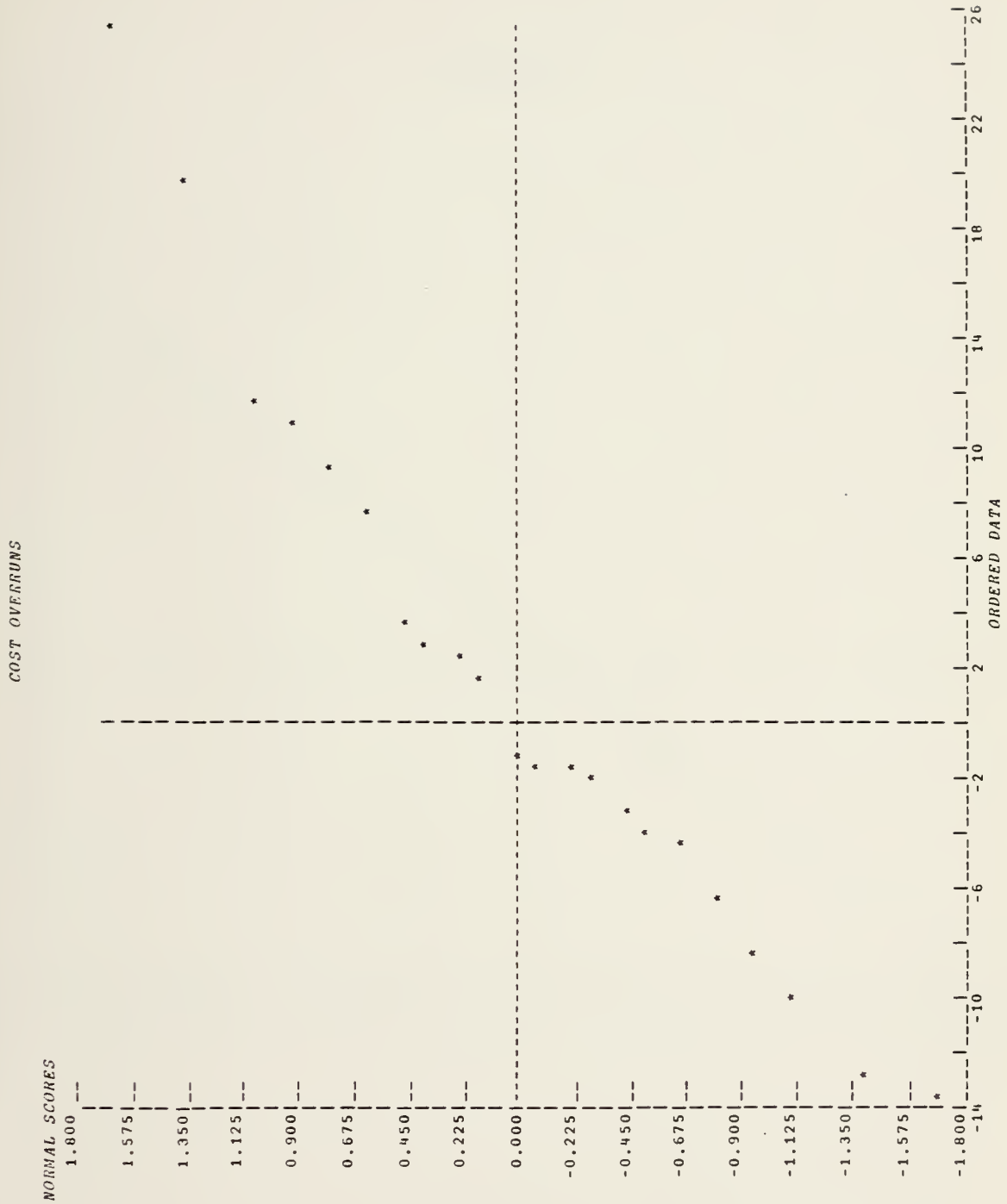
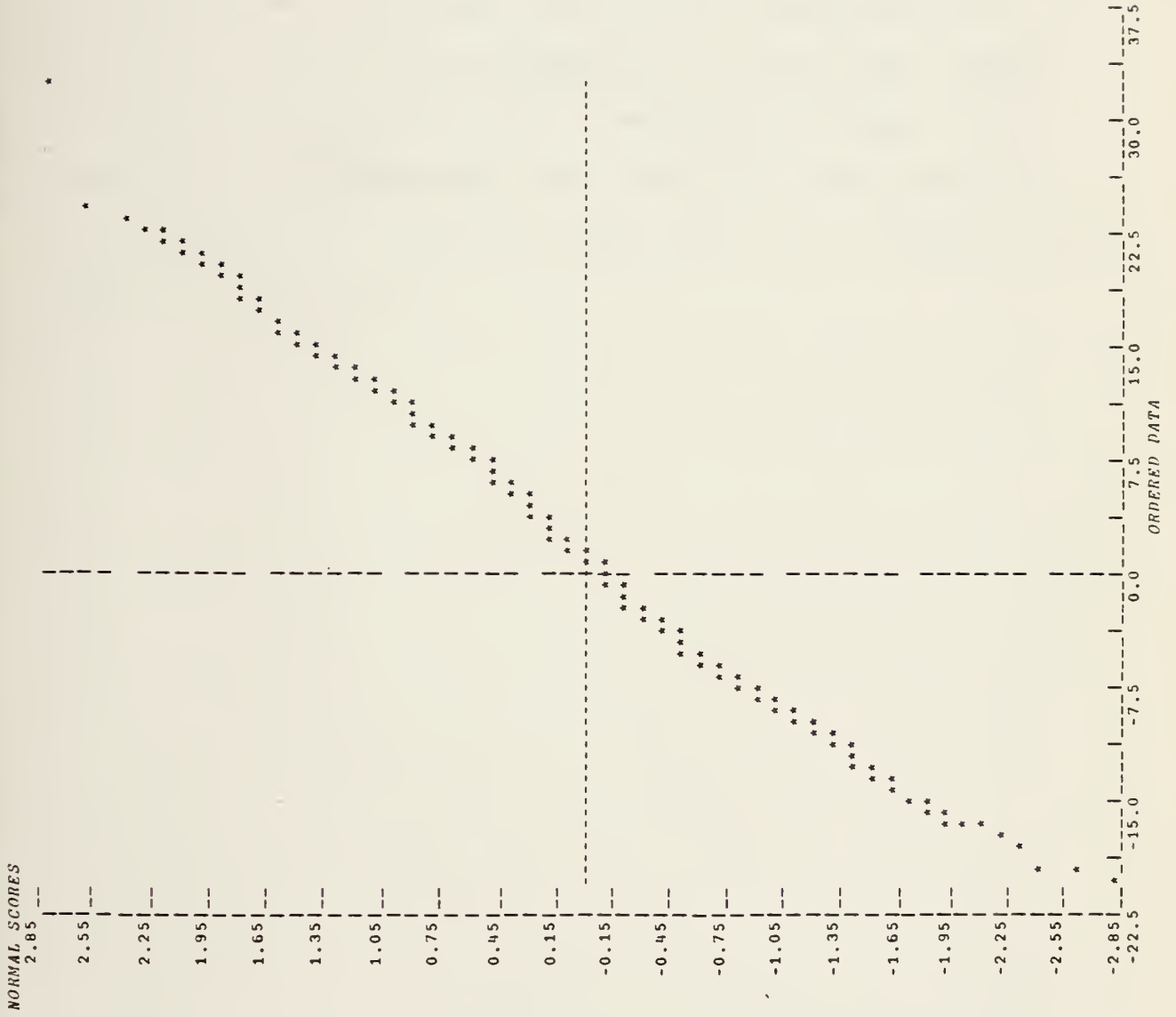


FIGURE 13

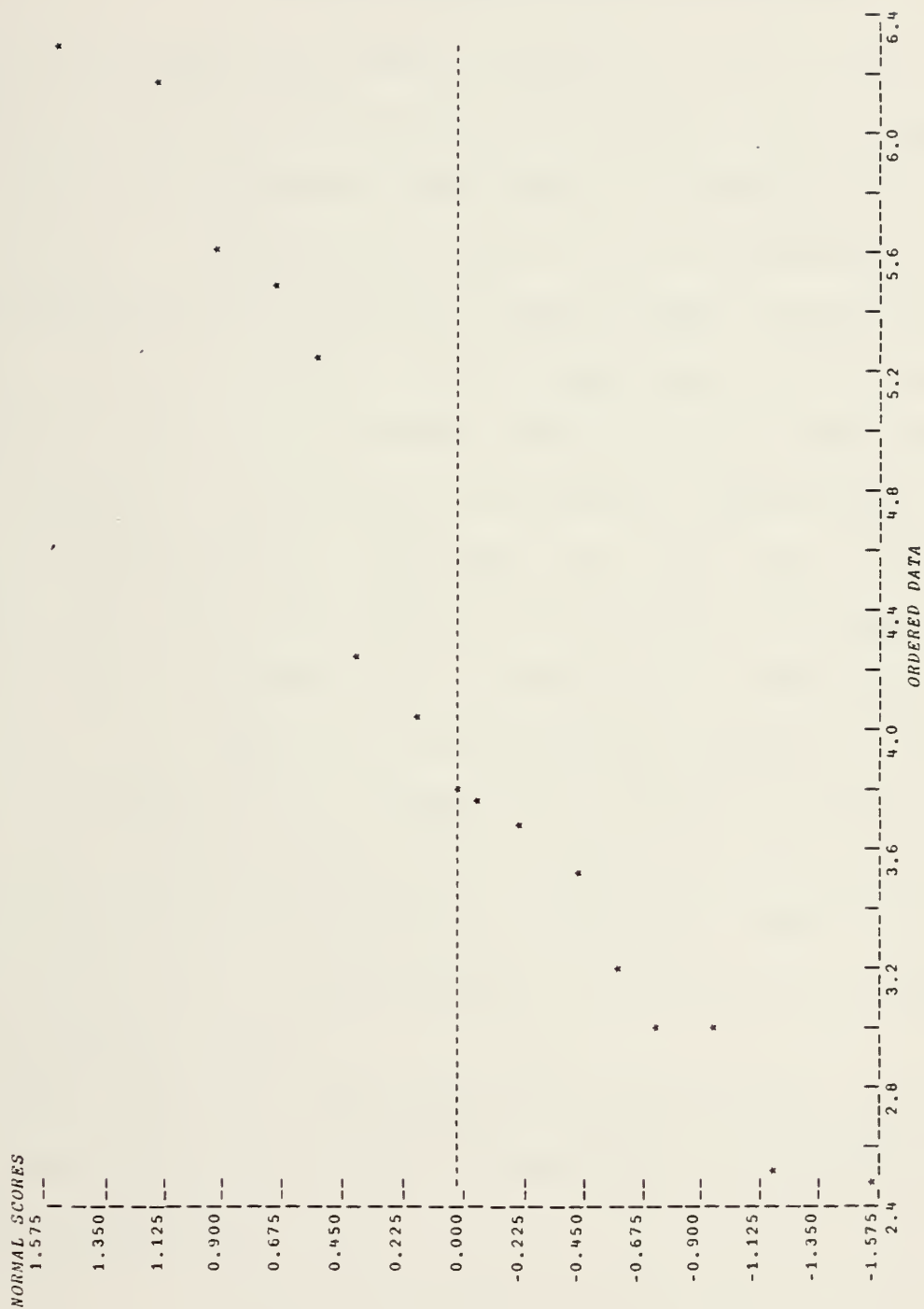
NORMAL SAMPLE GENERATED WITH SAME MEAN AND VARIANCE AS COST OVERRUN DATA



D. USAGE WITH COEFFICIENT OF VARIATION DATA OBTAINED
FROM USING SECTIONING ROUTINE

In order to check for normality in the sectioned estimates obtained from using HISTS (sectioning routine) the following was done. The 16 coefficient of variation values obtained from using HISTS with telephone data 1 (column 5, figure 7) were entered as a vector into NORMP . Figure 14 shows that the plot is marginally linear. This demonstrates the need for formal tests to verify normality in the absence of a strictly linear plot (Wilk & Gnanadsikan, 1968).

PLOT OF COEF VAR VALUES USING 16 SECTIONS FROM FIGURE 7



VIII. THE INDEPENDENCE AND MARKOV CHAIN HYPOTHESES FOR THE TELEPHONE DATA

The telephone data used in the thesis (Lewis & Cox, 1966) actually consists of binary bits transmitted over telephone lines and the information that the bit transmitted at time i , $i = 0, 1, 2, \dots$ is in error or not. This information is characterized by a sequence of binary-valued random variables $x(i)$, $i = 0, 1, \dots$ where $x(i)=1$ means that the bit transmitted at time i is in error, while $x(i)=0$ means that the bit transmitted at time zero is correctly transmitted.

In telephone data 1 there are 672 ones and 1,105,476 zeros, and a much more compact and equivalent representation of the data is obtained via the sequence of random variables $y(j)$, $j=1, 2, \dots$ where $y(j)$ is one plus the number of correctly transmitted bits between the j^{th} and $(j-1)^{\text{st}}$ bit error, with the convention that $y(j)=1$ if the errors occur on adjacent transmitted bits, and $y(1)$ is the time from $i=0$ to the first incorrectly transmitted bit. The $y(j)$ are called the times-between-errors.

A null hypothesis for the error structure which could be examined is that errors occur independently at each bit with a fixed probability, i.e.

$$P\{x(i)=1\} = \pi(1) \quad i=0, 1, \dots$$

$$P\{x(i)=0\} = \pi(0) = 1-\pi(1) \quad i=0, 1, \dots$$

The $y(j)$'s then are independent and geometrically distributed, since

$$P\{y(j)=1\} = P\{\text{if } (j-1)^{\text{st}} \text{ error at time } i; j^{\text{th}} \text{ at time } i+1\}$$

$$= \pi(1)$$

$$P\{y(j)=2\} = P\{\text{if } (j-1)^{\text{st}} \text{ error at time } i; j^{\text{th}} \text{ at time } i+2\}$$

$$= \pi(1)[1-\pi(1)] = \pi(1)\pi(0)$$

$$P\{y(j)=k+1\} = P\{\text{if } (j-1)^{\text{st}} \text{ error at time } i; j^{\text{th}} \text{ at time } i+1+k\}$$

$$= \pi(1)[1-\pi(1)]^k = \pi(1)[\pi(0)]^k$$

Note that, using the geometric series summation formula,

$$\sum_{k=1}^{\infty} P\{y(j)=k\} = \frac{\pi(1)}{1 - (1-\pi(1))} = 1$$

$$E[y(j)] = \sum_{k=1}^{\infty} kP\{y(j)=k\} = \frac{1}{1-\pi(0)} = \frac{1}{\pi(1)}$$

Now assume that the Markov structure of the zero's and ones is described by the transition matrix

$$\underline{P} = \begin{Bmatrix} P(0,0) & P(0,1) \\ P(1,0) & P(1,1) \end{Bmatrix} = \begin{Bmatrix} \rho + (1-\rho)\pi(1) & (1-\rho)\pi(0) \\ (1-\rho)\pi(1) & \rho + (1-\rho)\pi(0) \end{Bmatrix}$$

Here $P(m,n) = P\{x(i+1)=n \mid x(i)=m\}$, and we have parameterized the chain in terms of the stationary probability of a one or zero, and a correlation parameter $0 \leq \rho < 1$. Note that there are only two degrees of freedom in the stochastic

matrix, since rows must sum to 1, and there is only one degree of freedom if the stationary probability $\pi(0)=1-\pi(1)$ is fixed. Note that the stationary probabilities in the 2-state case are given by

$$\pi(0) = \frac{P(1,0)}{2-P(0,0)-P(1,1)} \quad \pi(1) = \frac{P(0,1)}{2-P(0,0)-P(1,1)}$$

We now define the runs of ones or zeros i.e. for $\ell=0$ or $\ell=1$, let

$$T_{\ell} = \inf\{n \geq 1: x(i+n) \neq \ell\} - 1,$$

the length of a run of ℓ 's, starting after time i , where the length can be $0, 1, 2, \dots$.

For example if $x(i+1)=1$, then the length of runs of zeros starting after time i is zero, the length of runs of ones is at least one long. Note that it is possible to talk of a conditional runs structure, i.e. the length of a run of ones which is given to start after time i . The run length is then at least one long.

Now the probability of a run T_{ℓ} having length greater than k is, using the Markov property,

$$P\{T_{\ell} \geq k\} = P\{x(i+1)=x(i+2)=\dots x(i+k)=\ell\} = \pi(\ell)[P(\ell, \ell)]^{k-1}$$

$k=1, \dots$

and $P\{T_{\ell} = 0\} = 1 - \pi(\ell).$

Thus, the run lengths are geometrically distributed and

$$E[T(\ell)] = \sum_{k=1}^{\infty} P\{T_{\ell} \geq k\} = \frac{\pi(\ell)}{1-P(\ell, \ell)} = \frac{\pi(\ell)}{(1-\rho)[1-\pi(\ell)]}$$

Note that $\rho=0$ gives the independence case, and while the runs of ones or zeros are geometrically distributed for both the independence or Markov dependent model, the mean run length is always longer for the Markov dependence, since

$$\frac{\pi(\ell)}{(1-\rho)[1-\pi(\ell)]} \geq \frac{\pi(\ell)}{[1-\pi(\ell)]} \quad 0 \leq \rho < 1$$

Thus, we could use the distributional properties of the runs to (1) check that either hypothesis is tenable or (2) if so, compare the estimated run lengths with the mean length $\hat{\pi}(\ell)/[1-\hat{\pi}(\ell)]$ predicted by the independence assumption. If the run lengths are not geometric, than another model must be postulated.

Note that when this mean time-between-errors is large as it is for telephone data 1 (figure 1; $E[y(j)] = 1,548$) the discreteness of the time scale can be ignored and the geometric distribution is indistinguishable from its continuous time analog, the exponential distribution.

That is approximation of the geometric distribution by an exponential distribution is valid can be seen from the fact that there are 672 errors ($x(i)$'s equal to one) in 1,106,148 transmitted bits, so that an estimate of $\pi(1)$, which is the maximum likelihood estimate under the independence hypothesis, is

$$\hat{\pi}(1) = \frac{\# x(i)'s = 1}{\text{total \# bits transmitted}} = \frac{\# x(i)'s = 1}{\# x(i)'s = 1 + \# x(i)'s = 0}$$

In the present data

$$\hat{\pi}(1) = \frac{672}{1,106,148} = .0006075$$

Now this geometric hypothesis will be examined, but it is clear from figure 1 that the hypothesis is not true. The distribution is in fact highly skewed and has been examined by Lewis & Cox, 1966.

An alternative model to independent bit errors is that the dependence structure is Markovian. One could examine this hypothesis with time-series methods but a method which is adaptable for use with the histogram routine and which examines both the independence and Markov assumptions is to look at runs of ones and zeros in the $x(i)$. Under both hypothesis these runs have geometrically distributed lengths.

The alternating conditional runs of ones for telephone data 1 are shown in figure 15 and for runs of zeros are shown in figure 16. Also, HISTLIST was used on the conditional runs and figure 17 shows the runs of ones and figure 18 shows the runs of zero.

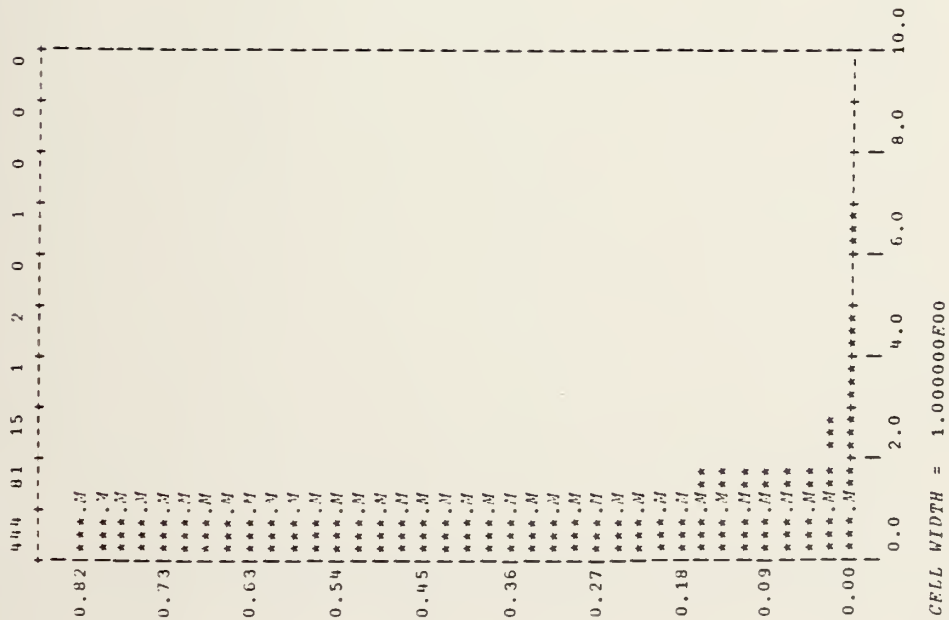
To test the hypothesis that the runs of ones in telephone data 1 is geometrically distributed the following was done.

Using figures 15 and 17 the following data was obtained:

MEAN	= 1.235294	# of runs = 1	= 444
VARIANCE	= .346008	# of runs = 2	= 81
		# of runs = 3	= 15
		# of runs = 4	= 1
		# of runs = 5	= 2
		# of runs ≥ 6	= 1

FIGURE 15

RUNS OF ONE FOR TELEPHONE DATA 1
FREQUENCIES
SAMPLE SIZE = 500



CENTRAL TENDENCY		SPREAD		HIGHER CENTRAL MOMENTS		DISTRIBUTION	
MEAN	1.235294E00	VARIANCE	3.460080E-01	M3	7.990531E-01	MINIMUM	1.000000E00
MEDIAN	1.000000E00	STD DEV	5.882245E-01	M4	3.219155E00	.10 QUANTILE (HINGE)	1.000000E00
TRIMEAN	1.000000E00	COEF VAR	4.761817E-01	SKEWNESS	3.925965E00	.25 QUANTILE (HINGE)	1.000000E00
MIDMEAN	1.000000E00	MEAN DEV	2.352941E-01	KURTOSIS	2.388688E01	.50 QUANTILE (MEDIAN)	1.000000E00
MIDRANGE	4.000000E00	RANGE	6.000000E00	BETA1	7.946519E-01	.75 QUANTILE (HINGE)	1.000000E00
GEOM MEAN	1.156667E00	MIDSPREAD	0.000000E00	BETA2	3.196802E00	.90 QUANTILE	2.000000E00
HARM MEAN	1.109541E00					MAXIMUM	7.000000E00



CELL WIDTH = 7.142857E02

CENTRAL TENDENCY		SPREAD		HIGHER CENTRAL MOMENTS		DISTRIBUTION	
MEAN	1.911270E03	VARIANCE	5.906497E07	M3	2.911020E12	MINIMUM	1.000000F00
MEDIAN	2.400000E01	STD DEV	7.685374E03	M4	1.855145E17	.10 QUANTILE	1.000000F00
TRIMEAN	5.087500E01	COEF VAR	4.021082E00	SKWNESS	6.412838F00	.25 QUANTILE (HINSE)	6.000000E00
MIDMEAN	3.894301E01	MEAN DEV	1.903406E03	KURTOSIS	5.017627E01	.50 QUANTILE (MEDIAN)	2.400000E01
RANGE	4.299650E04	RANGE	8.599100E04	BETA1	2.894986E12	.75 QUANTILE (HINSE)	1.495000F02
GEOM MEAN	3.822483E01	MIDSPREAD	1.435000E02	BETA2	1.894188E17	.90 QUANTILE	1.546000E03
HARM MEAN	5.585516E00					MAXIMUM	8.599200F04

HISTLIST
 HISTLIST PRINTS THE SERIAL NUMBER OF THE COMPRESSED
 DATA, THE ORDERED DATA COMPRESSED, AND THE NUMBER OF
 LIKE OCCURRENCES. ENTER YOUR DATA IN VECTOR FORM.
 []:
 ONE
 IF YOU WANT TO TITLE YOUR DATA TYPE YOUR TITLE.
 IF YOU DO NOT WANT A TITLE JUST HIT THE CARRIAGE
 RETURN.

RUNS OF ONE

IF YOU WANT YOUR OUTPUT TO APPEAR ON THE OFFLINE
 PRINTER TYPE 1 . IF YOU WANT YOUR OUTPUT TO APPEAR
 ON YOUR TERMINAL TYPE 0 .
 []:

0

RUNS OF ONE

SERIAL NUMBER	ORDERED DATA	NUMBER OF OCCURRENCES	PER CENT
1	1.000000	444	0.816
445	2.000000	81	0.149
526	3.000000	15	0.028
541	4.000000	1	0.002
542	5.000000	2	0.004
544	7.000000	1	0.002

FIGURE 17

FIGURE 18A

RUNS OF ZERO

SERIAL NUMBER ORDERED DATA NUMBER OF OCCURENCES PER CENT

1	1.000000	54	*****	0.099
2	2.000000	29	***	0.051
3	3.000000	22	**	0.040
4	4.000000	17	*	0.031
5	5.000000	11	*	0.020
6	6.000000	10	*	0.019
7	7.000000	12	*	0.022
8	8.000000	14	*	0.026
9	9.000000	9	*	0.017
10	10.000000	10	*	0.019
11	11.000000	11	*	0.020
12	12.000000	6	*	0.011
13	13.000000	6	*	0.011
14	14.000000	6	*	0.011
15	15.000000	8	*	0.015
16	16.000000	8	*	0.015
17	17.000000	5	*	0.009
18	18.000000	12	*	0.022
19	19.000000	1	*	0.002
20	20.000000	5	*	0.009
21	21.000000	5	*	0.009
22	22.000000	3	*	0.006
23	23.000000	7	*	0.013
24	24.000000	3	*	0.006
25	25.000000	3	*	0.006
26	26.000000	2	*	0.004
27	27.000000	3	*	0.006
28	28.000000	3	*	0.006
29	29.000000	6	*	0.011
30	30.000000	4	*	0.007
31	31.000000	4	*	0.007
32	32.000000	2	*	0.004
33	33.000000	4	*	0.007
34	34.000000	3	*	0.006
35	35.000000	2	*	0.004
36	36.000000	2	*	0.004
37	37.000000	1	*	0.002
38	38.000000	1	*	0.002
39	39.000000	2	*	0.004
40	40.000000	2	*	0.004
41	41.000000	1	*	0.002
42	42.000000	1	*	0.002
43	43.000000	4	*	0.007
44	44.000000	3	*	0.006
45	45.000000	1	*	0.002
46	46.000000	2	*	0.004
47	47.000000	1	*	0.002
48	48.000000	1	*	0.002
49	49.000000	1	*	0.002
50	50.000000	2	*	0.004
51	51.000000	1	*	0.002
52	52.000000	1	*	0.002
53	53.000000	3	*	0.006
54	54.000000	2	*	0.004
55	55.000000	2	*	0.004
56	56.000000	2	*	0.004
57	57.000000	1	*	0.002
58	58.000000	2	*	0.004
59	59.000000	1	*	0.002
60	60.000000	2	*	0.004
61	61.000000	1	*	0.002
62	62.000000	1	*	0.002
63	63.000000	1	*	0.002
64	64.000000	1	*	0.002
65	65.000000	1	*	0.002
66	66.000000	1	*	0.002
67	67.000000	1	*	0.002
68	68.000000	1	*	0.002
69	69.000000	1	*	0.002
70	70.000000	1	*	0.002
71	71.000000	1	*	0.002
72	72.000000	1	*	0.002
73	73.000000	1	*	0.002
74	74.000000	1	*	0.002
75	75.000000	1	*	0.002
76	76.000000	1	*	0.002
77	77.000000	1	*	0.002
78	78.000000	1	*	0.002
79	79.000000	1	*	0.002
80	80.000000	1	*	0.002
81	81.000000	1	*	0.002
82	82.000000	1	*	0.002
83	83.000000	1	*	0.002
84	84.000000	1	*	0.002
85	85.000000	1	*	0.002
86	86.000000	1	*	0.002
87	87.000000	1	*	0.002
88	88.000000	1	*	0.002
89	89.000000	1	*	0.002
90	90.000000	1	*	0.002
91	91.000000	1	*	0.002
92	92.000000	1	*	0.002
93	93.000000	1	*	0.002
94	94.000000	1	*	0.002
95	95.000000	1	*	0.002
96	96.000000	1	*	0.002
97	97.000000	1	*	0.002
98	98.000000	1	*	0.002
99	99.000000	1	*	0.002
100	100.000000	1	*	0.002

FIGURE 18B

3318	1118	.CCCCC00	1	0.002
3319	1118	.CCCCC00	1	0.002
3320	1119	.CCCCC00	4	0.007
3321	1120	.CCCCC00	2	0.004
3322	1121	.CCCCC00	2	0.004
3323	1122	.CCCCC00	4	0.007
3324	1123	.CCCCC00	3	0.006
3325	1124	.CCCCC00	1	0.002
3326	1125	.CCCCC00	1	0.002
3327	1126	.CCCCC00	1	0.002
3328	1127	.CCCCC00	1	0.002
3329	1128	.CCCCC00	1	0.002
3330	1129	.CCCCC00	1	0.002
3331	1130	.CCCCC00	1	0.002
3332	1131	.CCCCC00	1	0.002
3333	1132	.CCCCC00	1	0.002
3334	1133	.CCCCC00	1	0.002
3335	1134	.CCCCC00	1	0.002
3336	1135	.CCCCC00	1	0.002
3337	1136	.CCCCC00	1	0.002
3338	1137	.CCCCC00	1	0.002
3339	1138	.CCCCC00	1	0.002
3340	1139	.CCCCC00	1	0.002
3341	1140	.CCCCC00	1	0.002
3342	1141	.CCCCC00	1	0.002
3343	1142	.CCCCC00	1	0.002
3344	1143	.CCCCC00	1	0.002
3345	1144	.CCCCC00	1	0.002
3346	1145	.CCCCC00	1	0.002
3347	1146	.CCCCC00	1	0.002
3348	1147	.CCCCC00	1	0.002
3349	1148	.CCCCC00	1	0.002
3350	1149	.CCCCC00	1	0.002
3351	1150	.CCCCC00	1	0.002
3352	1151	.CCCCC00	1	0.002
3353	1152	.CCCCC00	1	0.002
3354	1153	.CCCCC00	1	0.002
3355	1154	.CCCCC00	1	0.002
3356	1155	.CCCCC00	1	0.002
3357	1156	.CCCCC00	1	0.002
3358	1157	.CCCCC00	1	0.002
3359	1158	.CCCCC00	1	0.002
3360	1159	.CCCCC00	1	0.002
3361	1160	.CCCCC00	1	0.002
3362	1161	.CCCCC00	1	0.002
3363	1162	.CCCCC00	1	0.002
3364	1163	.CCCCC00	1	0.002
3365	1164	.CCCCC00	1	0.002
3366	1165	.CCCCC00	1	0.002
3367	1166	.CCCCC00	1	0.002
3368	1167	.CCCCC00	1	0.002
3369	1168	.CCCCC00	1	0.002
3370	1169	.CCCCC00	1	0.002
3371	1170	.CCCCC00	1	0.002
3372	1171	.CCCCC00	1	0.002
3373	1172	.CCCCC00	1	0.002
3374	1173	.CCCCC00	1	0.002
3375	1174	.CCCCC00	1	0.002
3376	1175	.CCCCC00	1	0.002
3377	1176	.CCCCC00	1	0.002
3378	1177	.CCCCC00	1	0.002
3379	1178	.CCCCC00	1	0.002
3380	1179	.CCCCC00	1	0.002
3381	1180	.CCCCC00	1	0.002
3382	1181	.CCCCC00	1	0.002
3383	1182	.CCCCC00	1	0.002
3384	1183	.CCCCC00	1	0.002
3385	1184	.CCCCC00	1	0.002
3386	1185	.CCCCC00	1	0.002
3387	1186	.CCCCC00	1	0.002
3388	1187	.CCCCC00	1	0.002
3389	1188	.CCCCC00	1	0.002
3390	1189	.CCCCC00	1	0.002
3391	1190	.CCCCC00	1	0.002
3392	1191	.CCCCC00	1	0.002
3393	1192	.CCCCC00	1	0.002
3394	1193	.CCCCC00	1	0.002
3395	1194	.CCCCC00	1	0.002
3396	1195	.CCCCC00	1	0.002
3397	1196	.CCCCC00	1	0.002
3398	1197	.CCCCC00	1	0.002
3399	1198	.CCCCC00	1	0.002
3400	1199	.CCCCC00	1	0.002

FIGURE 18C

457	33000000	1	0.002
458	33000000	1	0.002
459	33000000	1	0.002
500	33000000	1	0.002
501	41000000	1	0.002
502	45000000	1	0.002
503	62000000	1	0.002
504	76000000	1	0.002
505	83000000	1	0.002
506	90000000	1	0.002
507	96000000	1	0.002
508	98000000	1	0.002
509	98000000	1	0.002
510	10000000	1	0.002
511	10000000	1	0.002
512	10000000	1	0.002
513	10000000	1	0.002
514	11000000	1	0.002
515	13000000	1	0.002
516	14000000	1	0.002
517	15000000	1	0.002
518	15000000	1	0.002
519	15000000	1	0.002
520	15000000	1	0.002
521	15000000	1	0.002
522	16000000	1	0.002
523	16000000	1	0.002
524	16000000	1	0.002
525	16000000	1	0.002
526	16000000	1	0.002
527	17000000	1	0.002
528	17000000	1	0.002
529	18000000	1	0.002
530	18000000	1	0.002
531	19000000	1	0.002
532	21000000	1	0.002
533	23000000	1	0.002
534	24000000	1	0.002
535	26000000	1	0.002
536	30000000	1	0.002
537	35000000	1	0.002
538	38000000	1	0.002
539	40000000	1	0.002
540	47000000	1	0.002
541	47000000	1	0.002
542	51000000	1	0.002
543	69000000	1	0.002
544	85000000	1	0.002

If the runs of ones are geometric then $\text{prob}\{x(i)=k\} = (1-p)p^{k-1}$ $k=1,2,\dots$. Thus, this is the "geometric plus one" distribution.

$$\mu = E[X] = \frac{1}{(1-p)}$$

$$\sigma^2 = \text{VAR}[X] = \frac{1}{(1-p)^2}$$

$$C(X) = \frac{\text{VAR}[X]^{\frac{1}{2}}}{E[X]} = p^{\frac{1}{2}}$$

To find p set $E[X] = 1.235294 = 1/(1-p)$

$$p = \underline{.1904761}$$

Therefore, if the data is "geometric plus one" then

$$\begin{aligned} \text{EXPECTED VAR}[X] &= .1904761/ (.8095329)^2 \\ &= \underline{.2906572} \end{aligned}$$

Thus, the expected variance is .2906572 and the observed variance from HIST is .3460080. Also, the expected coefficient of variance is

$$\text{EXPECTED } C(X) = (.1904761)^{\frac{1}{2}} = \underline{.4364356}$$

And, the observed coefficient of variation is .4761817.

Therefore, at this point there seems to be a fairly close agreement between the runs of one and a "geometric plus one" distribution with $p = .1904761$.

As further proof a Chi-square test for goodness of fit was run on the runs. By using the formula

$$\text{prob}\{X = x\} = (1-p)p^{x-1} \quad \text{for } x=1,2,3,4,5,\dots$$

<u>PROBABILITY</u>	<u>EXPECTED</u>	<u>OBSERVED</u>
P (X=1) = .8095239	440.38	444
P (X=2) = .1541949	83.88	81
P (X=3) = .0293704	15.98	15
P (X=4) = .0055943	3.04	1
P (X=5) = .0010655	.58	2
P (X $\geq$ 6) = .0002510	.14	1
	19.74	19

Note, to use Chi-square not more than 20% of the cells should have expected frequencies less than 5 and no cell should have an expected frequency less than one. Therefore, the above frequencies must be combined into 3 cells.

$$\chi^2 = \sum_{i=1}^3 \frac{(\text{obs}_i - \text{ex}_i)^2}{\text{ex}_i} = \underline{.1562799}$$

And, $\chi^2_{.05,2} = 5.99$. Thus, the null hypothesis that the runs of one are "geometric plus one" with $p = .1904761$ can not be rejected.

A similar procedure was done with the runs of greater than one. By using figure 15 the following information can be obtained:

MEAN = 1911.27
 VARIANCE = 59,064,970
 COEF.VAR.= 4.021082

And, by using the same method as previously done and solving for p one gets $p = \underline{.9994767}$.

$$\text{EXPECTED VAR}[X] = .9994767/((.0005233)^2) = \underline{3,651,213}$$

This expected variance differs greatly from the observed variance. Also, the expected coefficient of variation is

computed to be

$$\text{EXPECTED } C(X) = (.9994767)^{\frac{1}{2}} = \underline{.9997383}$$

This compares with the observed coefficient of variation of 4.021082 . Because of the gross departures of the variance and the coefficient of variation in the geometric hypothesis, one can conclude that the runs of length greater than 1 are not geometrically distributed.

IX. DOCUMENTATION ON ROUTINES

A. LOCATION IN APL LIBRARY

The descriptions and routines that have been presented are all available in the APL workspace library 2 DATALFNS . Providing the user is properly logged on the terminal and in the APL mode, all that is necessary is to type)LOAD 2 DATALFNS . If the user then types DESCRIBE, a short description of the six routines presented and instructions on how to obtain the detailed information that is available in each of the "HOW" variables would be printed.

B. WORKSPACE LOADING PROCEDURES

Each of the routines was designed to stand alone. That is, if the user desires just to use HIST , all that is necessary is to type)COPY 2 DATALFNS HISTGRP into a clear workspace. HISTGRP contains the principal routine HIST and only the additional routines necessary for HIST to operate. Thus, the user does not clutter his workspace with any unneeded functions. It is this group structure that maintains the orderliness of the workspace. And, the ability to copy a particular group into a clear workspace provides more space for data and executions of the functions.

The following is the group structure in library 2 DATALFNS .

<u>GROUP</u>	<u>PRINCIPAL ROUTINE</u>	<u>OTHER NECESSARY ROUTINES</u>	<u>VARIABLES</u>
HISTGRP	HIST	APLNAME,APLOT,AUTOS, CMS,DFT,ECDF,ECODE, EFT,OF,OUT,WRITE	
HISTLISTGRP	HISTLIST	APLNAME,CMS,ECODE, DFT,OF,OUT,WRITE	
HISTSGRP	HISTS	DFT,EFT	
HISTJACKGRP	HISTJACK	DFT,EFT,TOT	
EXPONPGRP	EXPONP	AND,AUTOSCALE, INITIAL,MPLLOT,MSGs, VS,MULTILOT,SETΔAP, TICMARK	<u>BS</u>
NORMPGRP	NORMP	AND,AUTOSCALE, INITIAL,MPLLOT,MSGs, VS,MULTILOT,SETΔAP, TICMARK	<u>BS</u>
DESCGRP (Descriptive group)			DESCRIBE,HISTHOW HISTHOW,HISTLIST- HOW,HISTJACKHOW, EXPONPHOW,NORMPHOW
VARIGRP (Variable group)			TELDAT1,TELDAT2, YROVR

C. ROUTINE LISTING

The above mentioned routines were either created by the author, adapted from existing fortran routine HISTG/F , or borrowed from the current APL library to supplement the author created routines.

1. Author Created Routines

HISTLIST, HISTS, HISTJACK, EXPONP, NORMP, APLLOT,
AUTOS, OUT, TOT

2. Adapted from Fortran Library Routine HISTG/F

HIST, ECDF

3. Borrowed Routines to Supplement Author Created Routines

AND, APLNAME, AUTOSCALE, CMS, DFT, ECODE, EFT,
INITIAL, MPLOT, MSGS, MULTILOT, NDTRI, OF, SET Δ AP, TICMARK,
VS, WRITE

X. COMPUTER LISTING OF ALL ROUTINES

```

VHIST[0]V
V HIST;M;V;X;T;A;APLN;B;C;D;DELTA;F;HSCALE;M1;M2;N;OL;SUM;SUMB;TT;XLABEL;Z;A1;A2;A3;A4;A41;A42;B1;B2;B3;C;B4;

T1
'ENTER DATA IN VECTOR FORM'
X+[]
.
.
'IF YOU ALSO WANT A SMOOTHED EMPIRICAL DENSITY FUNCTION ENTER'
'A 1 . IF YOU DO NOT WANT IT ENTER A 0 .'
B+[]
.
'IF YOU WANT TO TITLE YOUR HISTOGRAM TYPE YOUR TITLE.'
'IF YOU DO NOT WANT A TITLE JUST HIT THE CARRIAGE RETURN.'
TT+[]
.
.
'IF YOU WANT TO SET THE NUMBER OF CELLS AND THE SCALE ENTER'
'FIRST THE NUMBER OF CELLS (AN INTEGER BETWEEN 10 AND 28)'
'FOLLOWED BY A SPACE AND THEN YOUR LEFT SCALE POINT FOLLOWED'
'BY A SPACE AND THEN YOUR RIGHT SCALE POINT. HOWEVER, IF YOU'
'WANT HIST TO AUTOMATICALLY SCALE ENTER 0 0 0 .'
A+[]
.
.
'GIVEN THAT YOU HAVE SET YOUR OWN SCALE, TO INCLUDE DATA'
'POINTS THAT MIGHT BE OUTSIDE YOUR SCALE LIMITS IN THE END'
'CELLS, TYPE 1 . IF YOU DESIGNATED AUTOSCALE ALSO, TYPE'
'1 . IF HOWEVER, YOU DO NOT WANT THE DATA OUTSIDE THE SCALE'
LIMITS INCLUDED IN THE HISTOGRAM, TYPE 0 .'
Z+[]
.
.
'IF YOU WANT YOUR OUTPUT TO APPEAR ON THE OFFLINE PRINTER,'
'TYPE 1 . IF YOU WANT YOUR OUTPUT TO APPEAR ON YOUR'
'TERMINAL, TYPE 0 . (NOTE IF YOU TYPED 0 BE SURE YOUR'
'TERMINALS CARRIAGE PAGE SETTING IS ON THE MAXIMUM WIDTH)'
OL+[]
D+ 37 -3 4
C+29p0
+(OL=0)/TABJ
APLN+APLNAME 'HIST APLPF P1 V'
CMS 'ERASE HIST APLPF'
(20+'1') WRITE APLN
D+ 55 -5 6
TABJ: +(Z=0)/TAB2A
X+(X>A[2])^(X<A[3]))/X
TAB2A: C[12]+(C[27]+X[6X])-(C[21]+(X+X[AX])[1]
+(A[1]+0^A[2]+0^A[3]+0)/TAB1A
AUTOS

```



```

[44] TAB1A:=(A[1]<10)/LAST
[45] +(A[1]>28)/LAST
[46] DELTA:=(HSCALE+A[3]-A[2]):A[1]
[47] LABEL+A[2],(A[2]+DELTA+A[1])
[48] P+/(1/A[1]),=(X-A[2])+DELTA
[49] F[1]+/(X≤LABEL[1])+P[1]
[50] F[A[1]]+/(X>LABEL[A[1]+1])+F[A[1]]
[51] C[9]+(C[8]+/(X-C[1])+(X/X)*N)*2)+(N+P[X]-1)*0.5
[52] C[2]+0.5*(X/((N+2),1+(N+2)
[53] C[11]+/(X-C[2]))*N
[54] C[22]+X[(N+10)]
[55] C[29]+N-C[28]+(N+4)
[56] M2+1+M1+(1-((4|N):2))
[57] C[23]+(X[C[28]+1]+(X[C[28]]*M1))+M2
[58] C[24]+C[2]
[59] C[25]+(X[C[29]]+(X[C[29]+1]*M1))+M2
[60] C[26]+X[(0.9*N)]
[61] C[13]+C[25]-C[23]
[62] C[5]+(C[27]+C[21])*0.5
[63] C[3]+0.25*(C[23]+C[24]+C[24]+C[25])
[64] C[15]+((X-C[1])*3)*N*((N-1)*(N-2))
[65] C[19]+/(X-C[1])*3)*N
[66] C[16]+/(X-C[1])*4*(3+N*(N-2))+((N-1)*(N-2)*(N-3))
[67] C[20]+/(X-C[1])*4)*N
[68] C[16]+C[16]-(C[8]*C[8]*3*(N-1)*(N+N-3))+((N*(N-2)*(N-3))
[69] C[18]+3+C[16]+(C[8]*C[8])
[70] C[29]+N+1-C[28]+2+(N+4)
[71] SUN+C[23]+C[25]
[72] SUMA+X[(C[28]-1)]
[73] SUNB+X[(C[29])]
[74] SUN+SUN+(SUNB-SUMA)
[75] C[4]+SUM+(3+C[29]-C[28])
[76] C[17]+C[15]+C[9]*3
[77] C[6]+C[7]+0
[78] SUMA+5
[79] SUNB+7
[80] +(X[1]≤0)/TAB
[81] C[6]+*((X/(X)))*N)
[82] C[7]+N+/(X)
[83] SUMA+SUNB+7
[84] TAB:=(N>300)/TAB
[85] +(N>C[14]+P[1]+((P)V)(N)≤1)/V+X[(V=N+/(V+X/X)*X)/TAB2
[86] TAB4:C[14]+0
[87] SUNB+6
[88] TAB2:C[10]+0

```



```

[89]  *([C11]<1E 30)/TAB3
[90]  C[10]+C[9]+C[1]
[91]  TAB3:T1+CENTRAL TENDENCY
[92]  A1+MEAN MEDIAN SPREAD
[93]  A2+VARIANCE STD DEV TRIMEAN MIDMEAN MIDRANGE GEOM MEAN HARM MEAN DISTRIBUTION
[94]  A3+M3 M4 SKEWNESS KURTOSIS BETA1 BETA2 MIDSPREAD MODE
[95]  A41+MINIMUM .10 QUANTILE .25 QUANTILE (HINGE) .50 QUANTILE (MEDIAN)
[96]  A42+.75 QUANTILE (HINGE) .90 QUANTILE MAXIMUM
[97]  A4+A41.A42
[98]  B1+ 13 7 EFT B1+((SUMA).1)PB1+C[1].C[2].C[3].C[4].C[5].C[6].C[7]
[99]  B2+ 13 7 EFT B2+((SUMB).1)PB2+C[8].C[9].C[10].C[11].C[12].C[13].C[14]
[100] B3+ 13 7 EFT B3+ 6 1 PB3+C[15].C[16].C[17].C[18].C[19].C[20]
[101] B4+ 13 7 EFT B4+ 7 1 PB4+C[21].C[22].C[23].C[24].C[25].C[26].C[27]
[102] A1+((SUMA).11)PA1
[103] A2+((SUMB).11)PA2
[104] A3+ 7 10 PA3
[105] A4+ 7 23 PA4
[106] C+ 7 4 PC+
[107] APL0T
[108] TAB5:OL OUT T+
[109] OL OUT 'CELL WIDTH = ', 13 7 EFT DELTA
[110] B3+ 7 13 PB3+B3.T
[111] OL OUT 2 7 PT
[112] OL OUT T1
[113] OL OUT T
[114] +(SUMA=5)/TAB6
[115] +(SUMB=6)/TAB7
[116] TAB8:OL OUT A1.B1.C.A2.B2.C.A3.B3.C.A4.B4
[117] +(OL=0)/O
[118] CMS 'FINIS HIST APLPF'
[119] CMS 'O PRINTCC HIST APLPF'
[120] +(O=ECODE)/EX1
[121] 'HISTOGRAM SENT TO PRINTER'
[122] CMS 'ERASE HIST APLPF'
[123] +O
[124] EX1:O.p[1]+PRINTING FAILED. TRY AGAIN OR SEE APL PROGRAMMER.
[125] TAB6:A1+ 7 11 PA1
[126] B1+ 7 13 PB1
[127] A1[6;]+A1[7;]+B1[6;]+B1[7;]+
[128] +(SUMB=7)/TAB8
[129] TAB7:A2+ 7 11 PA2
[130] B2+ 7 13 PB2
[131] A2[7;]+B2[7;]+
[132] +TAB8
[133] LAST:NUMBER OF CELLS GIVEN IS NOT PERMISSABLE

```



```

VHISTLIST[[]]V
V HISTLIST;A;B;C;D;E;F;G;I;J;K;N;O;S;TT;APLN;OL
'HISTLIST PRINTS THE SERIAL NUMBER OF THE COMPRESSED'
[1] 'DATA. THE ORDERED DATA COMPRESSED, AND THE NUMBER OF'
[2] 'LIKE OCCURENCES. ENTER YOUR DATA IN VECTOR FORM.'
[3] X+[]
[4] . ,
[5] . ,
[6] 'IF YOU WANT TO TITLE YOUR DATA TYPE YOUR TITLE.'
[7] 'IF YOU DO NOT WANT A TITLE JUST HIT THE CARRIAGE'
[8] 'RETURN.'
[9] . ,
[10] TT+[]
[11] . ,
[12] 'IF YOU WANT YOUR OUTPUT TO APPEAR ON THE OFFLINE'
[13] 'PRINTER TYPE 1 . IF YOU WANT YOUR OUTPUT TO APPEAR'
[14] 'ON YOUR TERMINAL TYPE 0 .'
[15] OL+[]
[16] +(OL=0)/TABA
[17] APLN+APLNAME 'HISTLIST APLPF P1 V'
[18] CMS 'ERASE HISTLIST APLPF'
[19] (20+'1') WRITE APLN
[20] TABA:+(TT=0)/TAB10
[21] OL OUT 2 7 p' ,
[22] OL OUT TT
[23] OL OUT 2 7 p' ,
[24] TAB10:X+X[AX]
[25] C+(pX)p1
[26] K+1
[27] E+1p1
[28] J+0
[29] F+10
[30] TAB2:J+J+1
[31] +(X[J]=X[J+1])/TAB1
[32] K+K+1
[33] E+E,(J+1)
[34] F+F,X[J]
[35] +((J+1)=pX)/TAB3
[36] +TAB2
[37] TAB1:C[K]+C[K]+1
[38] +((J+1)=pX)/TAB3
[39] +TAB2

```



```

[40] TAB3:F+F,X[J+1]
[41] A←SERIAL NUMBER      ORDERED DATA      NUMBER OF OCCURENCES
[42] B←(0.5+60×((I/C)÷(pX)))
[43] D←8f(B+2)
[44] B←2f(B-4)p'
[45] OL OUT A,B,PER CENT'
[46] OL OUT '
[47] J←0
[48] TAB4:J+J+1
[49] G←(0.5+60×I÷(C[J]÷(pX)))
[50] B←G+Dp'
[51] G←Gp'
[52] I←3 DFT I
[53] S←10 0 DFT E[J]
[54] O←16 6 DFT F[J]
[55] N←10 0 DFT C[J]
[56] OL OUT S,'
[57] →(J=K)/TAB5      'O,'      'N,'      'G,B,I
[58] →TAB4
[59] TAB5:→(OL=0)/O
[60] CMS 'FINIS HISTLIST APLPF'
[61] CMS 'O PRINTCC HISTLIST APLPF'
[62] →(O=ECODE)/EX1
[63] 'HISTLIST SENT TO PRINTER'
[64] CMS 'ERASE HISTLIST APLPF'
[65] →0
[66] EX1:0,p[]←'PRINTING FAILED. TRY AGAIN OR SEE APL PROGRAMMER.'

```



```

VHISTS[[]]V
[1] HISTS:X;P;SE:ARRAY;J;FN:A;B;C;D;E;F;G;H;I;K;SD:VAR;MED;MIN;MAX;SDS;STS;KURT;SKEW;CVAR;MEAN;VRS;MNS;SZ;N3;M4;N
[2] 'TYPE THE NUMBER OF SECTIONS YOU DESIRE ( INTEGER
[3] 'BETWEEN 2 AND 28 ) BE SURE TO PICK YOUR NUMBER OF
[4] 'SECTIONS SO AS TO MINIMIZE THE NUMBER OF DATA
[5] 'POINTS THAT WILL HAVE TO BE DISCARDED. (HISTS
[6] 'PLACES THE DATA INTO THE EQUAL NUMBER OF SECTIONS
[7] 'YOU INDICATE DISCARDING ANY DATA LEFT OVER )
[8] SE+[]
[9] '
[10] 'ENTER YOUR DATA TO BE SECTIONED IN VECTOR FORM
[11] X+[]
[12] '
[13] P+0
[14] SDS+VRS+MNS+STS+7*0
[15] TAB10:SZ+L(PX)*SE
[16] MEAN+VAR+SD+CVAR+SKEW+MED+MIN+KURT+MAX+(SE)*0
[17] ARRAY+(SE),(SZ)*PX
[18] J+0
[19] TAB3:J+J+1
[20] '
[21] ' (J>SE)/TAB2
[22] MAX[J]+/(ARRAY[J;])
[23] MIN[J]+/(ARRAY[J;])
[24] SD[J]+(VAR[J]+(+(ARRAY[J;]-MEAN[J;]+(+/ARRAY[J;])*(N+SZ)*2)+(N+SZ)-1)*0.5
[25] FN+FN[FN]
[26] MED[J]+0.5*(+/FN[(N+2),1+(N+2)])
[27] M3+M4+(SE)*0
[28] M3[J]+((+/((ARRAY[J;]-MEAN[J;])*3))*N))*((N-1)*(N-2))
[29] M4[J]+((+/((ARRAY[J;]-MEAN[J;])*4))*N))*((N-1)*(N-2))*((N-1)*(N-3))
[30] M4[J]+M4[J]+(VAR[J]*VAR[J])*3*(N-1)*(N+2))*((N-1)*(N-2))*((N-3))
[31] SKEW[J]+M3[J]*SD[J]*3
[32] KURT[J]+3*M4[J]+(VAR[J]*VAR[J])
[33] CVAR[J]+SD[J]+MEAN[J]
[34] 'TAB3
[35] TAB2:-(P=1)/TAB12
[36] ARRAY+MEAN,MED,VAR,SD,CVAR,SKEW,KURT
[37] '
[38] '
[39] '
[40] '
[41] '
[42] '
[43] '
[44] '
[45] '
[46] '
[47] '
[48] '
[49] '
[50] '
[51] '
[52] '
[53] '
[54] '
[55] '
[56] '
[57] '
[58] '
[59] '
[60] '
[61] '
[62] '
[63] '
[64] '
[65] '
[66] '
[67] '
[68] '
[69] '
[70] '
[71] '
[72] '
[73] '
[74] '
[75] '
[76] '
[77] '
[78] '
[79] '
[80] '
[81] '
[82] '
[83] '
[84] '
[85] '
[86] '
[87] '
[88] '
[89] '
[90] '
[91] '
[92] '
[93] '
[94] '
[95] '
[96] '
[97] '
[98] '
[99] '
[100] '
[101] '
[102] '
[103] '
[104] '
[105] '
[106] '
[107] '
[108] '
[109] '
[110] '
[111] '
[112] '
[113] '
[114] '
[115] '
[116] '
[117] '
[118] '
[119] '
[120] '
[121] '
[122] '
[123] '
[124] '
[125] '
[126] '
[127] '
[128] '
[129] '
[130] '
[131] '
[132] '
[133] '
[134] '
[135] '
[136] '
[137] '
[138] '
[139] '
[140] '
[141] '
[142] '
[143] '
[144] '
[145] '
[146] '
[147] '
[148] '
[149] '
[150] '
[151] '
[152] '
[153] '
[154] '
[155] '
[156] '
[157] '
[158] '
[159] '
[160] '
[161] '
[162] '
[163] '
[164] '
[165] '
[166] '
[167] '
[168] '
[169] '
[170] '
[171] '
[172] '
[173] '
[174] '
[175] '
[176] '
[177] '
[178] '
[179] '
[180] '
[181] '
[182] '
[183] '
[184] '
[185] '
[186] '
[187] '
[188] '
[189] '
[190] '
[191] '
[192] '
[193] '
[194] '
[195] '
[196] '
[197] '
[198] '
[199] '
[200] '
[201] '
[202] '
[203] '
[204] '
[205] '
[206] '
[207] '
[208] '
[209] '
[210] '
[211] '
[212] '
[213] '
[214] '
[215] '
[216] '
[217] '
[218] '
[219] '
[220] '
[221] '
[222] '
[223] '
[224] '
[225] '
[226] '
[227] '
[228] '
[229] '
[230] '
[231] '
[232] '
[233] '
[234] '
[235] '
[236] '
[237] '
[238] '
[239] '
[240] '
[241] '
[242] '
[243] '
[244] '
[245] '
[246] '
[247] '
[248] '
[249] '
[250] '
[251] '
[252] '
[253] '
[254] '
[255] '
[256] '
[257] '
[258] '
[259] '
[260] '
[261] '
[262] '
[263] '
[264] '
[265] '
[266] '
[267] '
[268] '
[269] '
[270] '
[271] '
[272] '
[273] '
[274] '
[275] '
[276] '
[277] '
[278] '
[279] '
[280] '
[281] '
[282] '
[283] '
[284] '
[285] '
[286] '
[287] '
[288] '
[289] '
[290] '
[291] '
[292] '
[293] '
[294] '
[295] '
[296] '
[297] '
[298] '
[299] '
[300] '
[301] '
[302] '
[303] '
[304] '
[305] '
[306] '
[307] '
[308] '
[309] '
[310] '
[311] '
[312] '
[313] '
[314] '
[315] '
[316] '
[317] '
[318] '
[319] '
[320] '
[321] '
[322] '
[323] '
[324] '
[325] '
[326] '
[327] '
[328] '
[329] '
[330] '
[331] '
[332] '
[333] '
[334] '
[335] '
[336] '
[337] '
[338] '
[339] '
[340] '
[341] '
[342] '
[343] '
[344] '
[345] '
[346] '
[347] '
[348] '
[349] '
[350] '
[351] '
[352] '
[353] '
[354] '
[355] '
[356] '
[357] '
[358] '
[359] '
[360] '
[361] '
[362] '
[363] '
[364] '
[365] '
[366] '
[367] '
[368] '
[369] '
[370] '
[371] '
[372] '
[373] '
[374] '
[375] '
[376] '
[377] '
[378] '
[379] '
[380] '
[381] '
[382] '
[383] '
[384] '
[385] '
[386] '
[387] '
[388] '
[389] '
[390] '
[391] '
[392] '
[393] '
[394] '
[395] '
[396] '
[397] '
[398] '
[399] '
[400] '
[401] '
[402] '
[403] '
[404] '
[405] '
[406] '
[407] '
[408] '
[409] '
[410] '
[411] '
[412] '
[413] '
[414] '
[415] '
[416] '
[417] '
[418] '
[419] '
[420] '
[421] '
[422] '
[423] '
[424] '
[425] '
[426] '
[427] '
[428] '
[429] '
[430] '
[431] '
[432] '
[433] '
[434] '
[435] '
[436] '
[437] '
[438] '
[439] '
[440] '
[441] '
[442] '
[443] '
[444] '
[445] '
[446] '
[447] '
[448] '
[449] '
[450] '
[451] '
[452] '
[453] '
[454] '
[455] '
[456] '
[457] '
[458] '
[459] '
[460] '
[461] '
[462] '
[463] '
[464] '
[465] '
[466] '
[467] '
[468] '
[469] '
[470] '
[471] '
[472] '
[473] '
[474] '
[475] '
[476] '
[477] '
[478] '
[479] '
[480] '
[481] '
[482] '
[483] '
[484] '
[485] '
[486] '
[487] '
[488] '
[489] '
[490] '
[491] '
[492] '
[493] '
[494] '
[495] '
[496] '
[497] '
[498] '
[499] '
[500] '
[501] '
[502] '
[503] '
[504] '
[505] '
[506] '
[507] '
[508] '
[509] '
[510] '
[511] '
[512] '
[513] '
[514] '
[515] '
[516] '
[517] '
[518] '
[519] '
[520] '
[521] '
[522] '
[523] '
[524] '
[525] '
[526] '
[527] '
[528] '
[529] '
[530] '
[531] '
[532] '
[533] '
[534] '
[535] '
[536] '
[537] '
[538] '
[539] '
[540] '
[541] '
[542] '
[543] '
[544] '
[545] '
[546] '
[547] '
[548] '
[549] '
[550] '
[551] '
[552] '
[553] '
[554] '
[555] '
[556] '
[557] '
[558] '
[559] '
[560] '
[561] '
[562] '
[563] '
[564] '
[565] '
[566] '
[567] '
[568] '
[569] '
[570] '
[571] '
[572] '
[573] '
[574] '
[575] '
[576] '
[577] '
[578] '
[579] '
[580] '
[581] '
[582] '
[583] '
[584] '
[585] '
[586] '
[587] '
[588] '
[589] '
[590] '
[591] '
[592] '
[593] '
[594] '
[595] '
[596] '
[597] '
[598] '
[599] '
[600] '
[601] '
[602] '
[603] '
[604] '
[605] '
[606] '
[607] '
[608] '
[609] '
[610] '
[611] '
[612] '
[613] '
[614] '
[615] '
[616] '
[617] '
[618] '
[619] '
[620] '
[621] '
[622] '
[623] '
[624] '
[625] '
[626] '
[627] '
[628] '
[629] '
[630] '
[631] '
[632] '
[633] '
[634] '
[635] '
[636] '
[637] '
[638] '
[639] '
[640] '
[641] '
[642] '
[643] '
[644] '
[645] '
[646] '
[647] '
[648] '
[649] '
[650] '
[651] '
[652] '
[653] '
[654] '
[655] '
[656] '
[657] '
[658] '
[659] '
[660] '
[661] '
[662] '
[663] '
[664] '
[665] '
[666] '
[667] '
[668] '
[669] '
[670] '
[671] '
[672] '
[673] '
[674] '
[675] '
[676] '
[677] '
[678] '
[679] '
[680] '
[681] '
[682] '
[683] '
[684] '
[685] '
[686] '
[687] '
[688] '
[689] '
[690] '
[691] '
[692] '
[693] '
[694] '
[695] '
[696] '
[697] '
[698] '
[699] '
[700] '
[701] '
[702] '
[703] '
[704] '
[705] '
[706] '
[707] '
[708] '
[709] '
[710] '
[711] '
[712] '
[713] '
[714] '
[715] '
[716] '
[717] '
[718] '
[719] '
[720] '
[721] '
[722] '
[723] '
[724] '
[725] '
[726] '
[727] '
[728] '
[729] '
[730] '
[731] '
[732] '
[733] '
[734] '
[735] '
[736] '
[737] '
[738] '
[739] '
[740] '
[741] '
[742] '
[743] '
[744] '
[745] '
[746] '
[747] '
[748] '
[749] '
[750] '
[751] '
[752] '
[753] '
[754] '
[755] '
[756] '
[757] '
[758] '
[759] '
[760] '
[761] '
[762] '
[763] '
[764] '
[765] '
[766] '
[767] '
[768] '
[769] '
[770] '
[771] '
[772] '
[773] '
[774] '
[775] '
[776] '
[777] '
[778] '
[779] '
[780] '
[781] '
[782] '
[783] '
[784] '
[785] '
[786] '
[787] '
[788] '
[789] '
[790] '
[791] '
[792] '
[793] '
[794] '
[795] '
[796] '
[797] '
[798] '
[799] '
[800] '
[801] '
[802] '
[803] '
[804] '
[805] '
[806] '
[807] '
[808] '
[809] '
[810] '
[811] '
[812] '
[813] '
[814] '
[815] '
[816] '
[817] '
[818] '
[819] '
[820] '
[821] '
[822] '
[823] '
[824] '
[825] '
[826] '
[827] '
[828] '
[829] '
[830] '
[831] '
[832] '
[833] '
[834] '
[835] '
[836] '
[837] '
[838] '
[839] '
[840] '
[841] '
[842] '
[843] '
[844] '
[845] '
[846] '
[847] '
[848] '
[849] '
[850] '
[851] '
[852] '
[853] '
[854] '
[855] '
[856] '
[857] '
[858] '
[859] '
[860] '
[861] '
[862] '
[863] '
[864] '
[865] '
[866] '
[867] '
[868] '
[869] '
[870] '
[871] '
[872] '
[873] '
[874] '
[875] '
[876] '
[877] '
[878] '
[879] '
[880] '
[881] '
[882] '
[883] '
[884] '
[885] '
[886] '
[887] '
[888] '
[889] '
[890] '
[891] '
[892] '
[893] '
[894] '
[895] '
[896] '
[897] '
[898] '
[899] '
[900] '
[901] '
[902] '
[903] '
[904] '
[905] '
[906] '
[907] '
[908] '
[909] '
[910] '
[911] '
[912] '
[913] '
[914] '
[915] '
[916] '
[917] '
[918] '
[919] '
[920] '
[921] '
[922] '
[923] '
[924] '
[925] '
[926] '
[927] '
[928] '
[929] '
[930] '
[931] '
[932] '
[933] '
[934] '
[935] '
[936] '
[937] '
[938] '
[939] '
[940] '
[941] '
[942] '
[943] '
[944] '
[945] '
[946] '
[947] '
[948] '
[949] '
[950] '
[951] '
[952] '
[953] '
[954] '
[955] '
[956] '
[957] '
[958] '
[959] '
[960] '
[961] '
[962] '
[963] '
[964] '
[965] '
[966] '
[967] '
[968] '
[969] '
[970] '
[971] '
[972] '
[973] '
[974] '
[975] '
[976] '
[977] '
[978] '
[979] '
[980] '
[981] '
[982] '
[983] '
[984] '
[985] '
[986] '
[987] '
[988] '
[989] '
[990] '
[991] '
[992] '
[993] '
[994] '
[995] '
[996] '
[997] '
[998] '
[999] '
[1000] '

```



```

[46]      ' '
[47]      TAB12:J+0
[48]      TAB6:J+J+1
[49]      K+ 2 0 DFT J
[50]      A+ 11 5 EFT MEAN[J]
[51]      B+ 11 5 EFT MED[J]
[52]      C+ 11 5 EFT VAR[J]
[53]      D+ 11 5 EFT SD[J]
[54]      E+ 11 5 EFT CVAR[J]
[55]      F+ 11 5 EFT SKEW[J]
[56]      G+ 11 5 EFT KURT[J]
[57]      H+ 11 5 EFT MIN[J]
[58]      I+ 11 5 EFT MAX[J]
[59]      + (P=1)/TAB7
[60]      ' ' K ' ' A ' ' B ' ' C ' ' D ' ' E ' ' F ' ' G ' ' H ' ' I
[61]      + (J=SE)/TAB11
[62]      +TAB6
[63]      TAB11:P+SE+1
[64]      +TAB10
[65]      TAB7: 2 7 0 ' '
[66]      'UNSECTIONED' ' ' A ' ' B ' ' C ' ' D ' ' E ' ' F ' ' G ' ' H ' ' I
[67]      2 1 0 ' '
[68]      'SUMMARY FOR SECTIONED DATA'
[69]      ' '
[70]      ' ' MEAN
[71]      ' '
[72]      J+0
[73]      TAB8:J+J+1
[74]      A+ 11 5 EFT MNS[J]
[75]      B+ 11 5 EFT VRS[J]
[76]      C+ 11 5 EFT SDS[J]
[77]      D+ 11 5 EFT STS[J]
[78]      E+MEAN MEDIAN
[79]      E+ 7 12 0E
[80]      E[J]:A ' ' B ' ' C ' ' D
[81]      + (J=7)/O
[82]      +TAB8

```

KURTOSIS

SKEWNESS

COEF VAR

STD DEV

VARIANCE

MEDIAN

STD DEV

COEF VAR

SKEWNESS

KURTOSIS


```

V HISTJACK[1]J
V HISTJACK;X;SEC1;PSV;SZ;A;B;C;J;G;ARRAY;BRRAY;K;FN;S;CVAR;KURT;MAX;MEAN;MEANS;MED;MIN;M3;M4;N;SD;SKEW;VAR;VARSA
*TYPE THE NUMBER OF GROUPS YOU DESIRE (INTEGER
BETWEEN 2 AND 50) BE SURE TO PICK YOUR NUMBER
OF GROUPS SO AS TO MINIMIZE THE NUMBER OF DATA
POINTS THAT WILL HAVE TO BE DISCARDED. (HISTJACK
PLACES THE DATA INTO THE EQUAL NUMBER OF GROUPS
YOU INDICATE DISCARDING ANY DATA LEFT OVER)'
SEC1+[]
[1] [2] SEC1+[]
[3] [3] ' '
[4] [4] 'ENTER YOUR DATA TO BE JACKKNIFED IN VECTOR FORM'
[5] [5] X+[]
[6] [6] ' '
[7] [7] MEANS+VARSA+S+ 7 1 p0
[8] [8] PSV+((7),(SEC1))p0
[9] [9] SZ+(pX)-(1((pX)+SEC1))
[10] [10] MEAN+VAR+SD+CVAR+SKEW+MED+MIN+KURT+MAX+(SEC1+1)p0
[11] [11] ARRAY+(1,(pX))pX
[12] [12] J+G+1
[13] [13] B+pX
[14] [14] TAB3:=(J>(SEC1+1))/TAB2
[15] [15] MAX[J]+/(ARRAY[G;])
[16] [16] MIN[J]+/(ARRAY[G;])
[17] [17] SD[J]+(VAR[J]+(+/(ARRAY[G;]-MEAN[J])*(N-1)*(N-2))+(N+B-1)*0.5
[18] [18] FN+ARRAY[G;]
[19] [19] FN+FN[4FN]
[20] [20] MED[J]+0.5*(+/(FN*(N+2),1+(N+2)))
[21] [21] M3+M4+(SEC1+1)p0
[22] [22] M3[J]+((+/(ARRAY[G;]-MEAN[J])*3))*N)+((N-1)*(N-2))
[23] [23] M4[J]+((+/(ARRAY[G;]-MEAN[J])*4))*N)+((N-1)*(N-2))+(N-3)
[24] [24] M4[J]+M4[J]-((VAR[J]*3*(N-1)*(N+2))+(N*(N-2)*(N-3)))
[25] [25] SKEW[J]+M3[J]*SD[J]*3
[26] [26] KURT[J]+3+M4[J]+(VAR[J]*VAR[J])
[27] [27] CVAR[J]+SD[J]*MEAN[J]
[28] [28] G+(J+J+1)-1
[29] [29] B+SZ
[30] [30] +(G22)/TAB3
[31] [31] C+1((pX)+A+SEC1
[32] [32] BRRAY+((A),(C))pX
[33] [33] ARRAY+((A),(S2))p0

```



```

[34] 101
[35] +TAB3
[36] TAB2:PSV[1,]+(A*MEAN[1])-(A-1)*MEAN[1+1SEC1])
[37] PSV[2,]+(A*MED[1])-(A-1)*MED[1+1SEC1])
[38] PSV[3,]+(A*VAR[1])-(A-1)*VAR[1+1SEC1])
[39] PSV[4,]+(A*SD[1])-(A-1)*SD[1+1SEC1])
[40] PSV[5,]+(A*CVAR[1])-(A-1)*CVAR[1+1SEC1])
[41] PSV[6,]+(A*SKEW[1])-(A-1)*SKEW[1+1SEC1])
[42] PSV[7,]+(A*KURT[1])-(A-1)*KURT[1+1SEC1])
[43] MEANS+((+/PSV)*A)
[44] VARSA+((+/PSV*2)-((+/PSV)*2)*SEC1))*((SEC1-1)
[45] S+(VARSA*SEC1)*0.5
[46] A+GROUP
[47] C+SKEWNESS KURTOSIS MEAN MEDIAN VARIANCE STD DEV COEF VAR
[48] , ,
[49] A,C
[50] , ,
[51] J+1
[52] A+ 9 11 0 , ,
[53] TAB4:J+J+1
[54] A[1,]+ 11 5 EFT MEAN[J]
[55] A[2,]+ 11 5 EFT MED[J]
[56] A[3,]+ 11 5 EFT VAR[J]
[57] A[4,]+ 11 5 EFT SD[J]
[58] A[5,]+ 11 5 EFT CVAR[J]
[59] A[6,]+ 11 5 EFT SKEW[J]
[60] A[7,]+ 11 5 EFT KURT[J]
[61] A[8,]+ 11 5 EFT MIN[J]
[62] A[9,]+ 11 5 EFT MAX[J]
[63] K+ 2 0 DFT(J-1)
[64] +(J=1)/TAB6
[65] ,K, ,A[1,], ,A[2,], ,A[3,], ,A[4,], ,A[5,], ,A[6,], ,A[7,], ,A[8,], ,A[9,]
[66] +(J=(SEC1+1))/TAB5
[67] +TAB4
[68] TAB5:J+0
[69] +TAB4

```



```

[70] TAB6: 2 1 p' '
[71] 'UNGROUPED' 'A[1:],' 'A[2:],' 'A[3:],' 'A[4:],' 'A[5:],' 'A[6:],' 'A[7:],' 'A[8:],' 'A[9:]
[72] 2 1 p' '
[73] 'SUMMARY FOR JACKKNIFE DATA'
[74] ' '
[75] ' ' JACKKNIFE ESTIMATE VARIANCE (VARGROUPS)*.5'
[76] A+48p' '
[77] A.' JACKKNIFE ESTIMATE OF STD DEV'
[78] A.' OF MEAN OF PSEUDO-VALUES'
[79] ' '
[80] A+'MEAN MEDIAN VARIANCE STD DEV COEF VAR SKEWNESS KURTOSIS '
[81] A+ 7 9 pA
[82] J+0
[83] C+ 3 11 p' '
[84] TAB7: J+J+1
[85] C[1:]+ 11 5 EFT MEANS[J]
[86] C[2:]+ 11 5 EFT VARSA[J]
[87] C[3:]+ 11 5 EFT S[J]
[88] A[J:],' 'C[1:],' 'C[2:],' 'C[3:]
[89] +(J=7)/TAB8
[90] +TAB7
[91] TAB8:+0

```



```

VEXPONP([ ])V
EXPONP;X;Y;TT;SM;EQ;BL;TL;EQ2;R90;UL;QL;ET;PI;ST;D
'EXPONP ORDERS THE DATA YOU GIVE AND COMPUTES THE
'EMPIRICAL LOG SURVIVER FUNCTION FOR THE DATA.'
'A PLOT OF THE LOG SURVIVER FUNCTION FOR THE DATA.'
'IS THEN PRINTED TO SEE IF THERE IS A LINEAR FIT.'
' '
'IF YOU WANT TO TITLE YOUR PLOT TYPE YOUR TITLE.'
'IF YOU DO NOT WANT A TITLE JUST HIT THE CARRIAGE'
'RETURN.'
' '
[10] TT+1
[11] '
[12] 'ENTER YOUR DATA IN VECTOR FORM'
[13] X+1
[14] SM+ 3 10
[15] '
[16] EQ+'*+O'
[17] BL+'ORDERED DATA'
[18] TL+'EXPONENTIAL SCORES'
[19] EQ2+'*'
[20] R90+H2+QL+O
[21] D+130
[22] ST+ 1 1.25 1.5 2 2.5 3 4 5 7.5 10
[23] PL+10
[24] BT+10
[25] X+X[AX]
[26] Y+((((P*1)-(P*1)))+(P*1))
[27] (10 10 ,(P*1)) MPLOT Y VS X
V

```

```

VNDTRI([ ])V
R+NDTRI P;W
[1] R+ 0 0
[2] +((0.5≤|P-0.5|,0=|P-0.5|)/L1,L2
[3] W+(P*(1-P)*2)*0.5
[4] L6:R[1]+(W-((+/(2.515517 0.802853 0.010328)*W*-1+1,3))+/(1 1.432788 0.189269 0.001308)*W*-1+1,4))*P-
0.5
[5] R[2]+0.3989423*-0.5*R[1]*R[1]
[6] +0
[7] L1:R[1]+(-P-0.5)*10*74
[8] +0
[9] L2:R[2]+0.3989423
V

```

```

VECDF([ ])V
ECDF;XN;BN;FN;LIN;I;FMA;J;U;LOW;XINC;MAXM1;TRY
XN+BN+((N)*0.5)+C[12])!W
[1] LOW+I+1
[2] FMA+10
[3] XINC+HSCALE:(A[1]*4)
[4] FN+(MAXM1+((A[1]*4)-1))*0
[5] TAB60:+(I>MAXM1)/TAB50
[6] U+A[2]+I*XINC
[7] J+LOW
[8] TAB61:+(J>N)/TAB53
[9] TRY+BN*(U-X[J])
[10] +(TRY>1)/TAB52
[11] +(TRY<1)/TAB53
[12] FN[I]+FN[I]+1-|TRY
[13] J+J+1
[14] +TAB61
[15] TAB52:LOW+J
[16] J+J+1
[17] +TAB61
[18] TAB53:FN[I]+FN[I]*XN
[19] FMA+FMA+FN[I]
[20] I+I+1
[21] +TAB60
[22] TAB50:I+1
[23] TAB62:+(I>MAXM1)/TAB54
[24] LIN+1((D[1]+0.5)-(D[1]-1)*(FN[I]+FMA))
[25] ARRAY[LIN;I]+*F
[26] I+I+1
[27] +TAB62
[28] TAB54:+0
[29]
V

```



```

V NORMP[[]]V
V NORMP:X;T;R;L;S;M;T;L;P;C2;I;J;R;R90;H;E;C;L;H;T;P;I;S;T;D;P;C
'NORM ORDERS THE DATA YOU GIVE AND COMPUTES THE'
'INVERSE OF THE UNIT NORMAL CUMULATIVE DISTRIBU-'
'TION FOR THE DATA. A PLOT OF THE INVERSE OF THE'
'UNIT NORMAL CUMULATIVE DISTRIBUTION VS THE ORDER-'
'ED DATA IS THEN PRINTED TO SEE IF THERE IS A'
'LINEAR FIT.'
'IF YOU WANT TO TITLE YOUR PLOT TYPE YOUR TITLE.'
'IF YOU DO NOT WANT A TITLE JUST HIT THE CARRIAGE'
'RETURN.'
T;T+V
'ENTER YOUR DATA IN VECTOR FORM'
X+[]
'ORDERED DATA'
S;M+ 3 10
R90+H;S+G;L+0
D+130
S;T+ 1 1.25 1.5 2 2.5 3 4 5 7.5 10
P;I+10
B;T+10
T;L+ 'NORMAL SCORES'
P;C2+ ''
P;C+ '*-|'
X+X[ΔX]
R+((ρX),2)ρ0
I+((1(ρX)):(ρX)+1)
J+0
TAB3:J+J+1
R[J:]←NDTRI I[J]
+((J=(ρX))/TAB2
→TAB3
TAB2:J+((R[;1]),101ρ0
I+X,X[1],(X[1]+((X[(ρX)]-X[1]):100)×100)
+((X[1]<0ΔX[(ρX)]>0)/TAB4
(10 10 ,(ρX),(101)) MPLOT J VS I
→TAB5
TAB4:J+J,P[1;1],(R[1;1]+((R[(ρX);1]-R[1;1]):48)×148)
I+I,49ρ0
(10 10 ,(ρX),(101),(49)) MPLOT J VS I
TAB5:→0
V
V AUTOS[[]]V
V AUTOS
[1] →((ρX)≥80)/TAB70
[2] A[1]+16
[3] →TAB71
[4] TAB70:A[1]+((28|(ρX+5)))
[5] TAB71:A[2]+C[21]
[6] A[3]+C[27]
[7] +0
V
VEFT[[]]V
V Z+W EFT X;D;E;H;J;K;L;Q;S;T;U;Y
P+ '0123456789.E -'
[1] →(V/W≠|W+,(H+0)×L+1<ρX)/EFTERR+0×K+2
[2] →(3 2 1 <ρX)/(EFTERR+K+0), 2 3 +I26
[3] X+((V/ 1 2 =ρW)φ 1 2)Q(1,ρ,X)ρX
[4] X+(φ2ρφX)ρX
[5] →((Δ/(ρW))≠ 1 2 ,2×E+1ρφX),1×ρW)/(EFTERR×K+1),2+I26
[6] W+(W+6+(V/X<0)+V/1>|X),W
[7] →(V/6>-/[1] W+Q(E,2)ρW)/EFTERR+0×K+2
[8] Z+((K+1ρX),+W[1;1])ρ'
[9] EFTLP:→(E<H+H+1)/EFTEND
[10] S+1+[10●|Y+0=Y+X[:H]
[11] U+1+[10●|Y+0=Y+10.5+(10×Q-15)+Y×10*(Q+W[2;H])-S
[12] J+((T-4)ρ1),4ρ0),1+[10|(|Y:10×U>Q)○.10×1+φ1 4+T+W[1;H]
[13] J[:T- 2 1]+1+[10|(|S-U≤Q)○.10 1
[14] J[:1,U,T,T-3]+Q(4,K)ρ(Kρ11),(13+0>Y,S-1),Kρ12
[15] J[:1,U,T,T-3]+J[:1φ1U+1),(U+1+1Q)]
[16] J[:T- 2 1 0]+(-S≤0)φJ[:T- 2 1 0]
[17] →EFTLP,ρZ[:(+W[1;1;H-1])+1,T]+D[J]
[18] EFTEND:→L/0
[19] +0×ρZ+,Z
[20] EFTERR: 'EFT',(3 6 ρ' RANK LENGTHDOMAIN')[K+1;], 'PROBLEM.'
[21]
[22]
V

```



```

VAPLOT(U)
V
  APLOT;I;J;LINE,CROB;PROB;VERT;H1;PLABEL;DIB;FSCALE;DID;DIT;DIS;IQ1T;IQ2T;IQ3T;NMAX;MNT;RT;PRBMX;INCR;ARRAY;FMAX
  +((P*TT)=(P*0))/TAB5A
  OL OUT(18P','),TT
  OL OUT 10','
  TAB5A:ARRAY+(D[1]),(4*A[1])P','
  FSCALE+(D[1]-1)*FMAX+(F[1]*F)[P*F]
  I+0
  J+3
  TAB12:+(I=A[1])/TAB15
  I+I+1
  J+J+4
  LINE+1((D[1]-0.5)-(FSCALE*F[I]))
  +((LINE>=D[1]-1))/TAB13
  ARRAY[LINE+(D[1]-LINE);J]+','
  ARRAY[LINE+(D[1]-LINE);J+1]+','
  ARRAY[LINE+(D[1]-LINE);J+2]+','
  ARRAY[LINE+(D[1]-LINE);J+3]+','
  ARRAY[(D[1]);J+3]+','
  +TAB12
  TAB13:+(F[I]≠0)/TAB14
  ARRAY[(D[1]);J]+','
  ARRAY[(D[1]);J+1]+','
  ARRAY[(D[1]);J+2]+','
  ARRAY[(D[1]);J+3]+','
  +TAB12
  TAB14:ARRAY[(D[1]);J]+','
  ARRAY[(D[1]);J+1]+','
  ARRAY[(D[1]);J+2]+','
  ARRAY[(D[1]);J+3]+','
  +TAB12
  TAB15:PROB+((D[1]),4)P','
  INCR+(PRBMX+FMAX+H)*9
  CROB+PRBMX,(PRBMX-INCR*18),0
  CROB+ 4 2 DFT CROB+ 10 1 0CROB
  PROB[D[2]+D[3]*10;]+CROB[110;]
  VERT+((D[1]),1)P','
  RT+(NMAX+A[1]*4)+(A[3]-A[2])
  IQ1T+(0.5+(C[23]-A[2])*RT)
  IQ2T+(0.5+(C[24]-A[2])*RT)
  IQ3T+(0.5+(C[25]-A[2])*RT)
  MNT+(0.5+(C[1]-A[2])*RT)
  +(MNT>1)/TAB21
  MNT+1
  TAB21:+(MNT≤NMAX)/TAB22
  MNT+NMAX
  TAB22:+(IQ1T>1)/TAB23

```



```

[46] IQ1T,1
[47] TAB23:->(IQ2T>1)/TAB24
[48] IQ2T+1
[49] TAB24:->(IQ3T>1)/TAB25
[50] IQ3T+1
[51] TAB25:->(IQ1T<=NMAX)/TAB26
[52] IQ1T+NMAX
[53] TAB26:->(IQ2T<=NMAX)/TAB27
[54] IQ2T+NMAX
[55] TAB27:->(IQ3T<=NMAX)/TAB28
[56] IQ3T+NMAX
[57] TAB28:ARRAY[1(D[1]);IQ1T]+,.,
[58] ARRAY[1(D[1]);IQ2T]+,.,
[59] ARRAY[1(D[1]);IQ3T]+,.,
[60] ARRAY[1(D[1]);MNT]+,M,
[61] ARRAY[1(D[1]);NMAX]+,|,
[62] +(B=0)/TAB3A
[63] ECDF
[64] TAB3A:H1+(5p',),H1+,'FREQUENCIES',H1+(32p',),H1+,'SAMPLE SIZE = ',
[65] OL OUT H1, 6 0 DFT(pX)
[66] OL OUT 1p',
[67] OL OUT H1+(4p',),H1+ 4 0 DFT F
[68] OL OUT ',+',(4xA[1])p',---+',
[69] OL OUT ',|',H1+((4*(A[1]-1))p',),H1+ '|',
[70] OL OUT PROB,VERT,ARRAY
[71] DIS+1((pXLABEL)+2)
[72] DIT+(pXLABEL)+2
[73] +((DIT-DIS)=0)/TAB40
[74] TAB41:DID+(8*DIS)p',|
[75] OL OUT ',DID,|',
[76] +TAB42
[77] TAB40:DIS+DIS-1
[78] +TAB41
[79] TAB42:DIB+DIS+1
[80] XLABEL+XLABEL['_1+2*1,DIB]
[81] +(XLABEL>9999)/TAB31
[82] +(XLABEL<9999)/TAB31
[83] +(DELTA<0.1)/TAB31
[84] PLABEL+((pXLABEL),1)pXLABEL
[85] OL OUT ',PLABEL+,PLABEL+(PLABEL+ 7 1 DFT PLABEL),((pXLABEL),1)p',
[86] +0
[87] TAB31:XLABEL+XLABEL['_1+2*1(|((pXLABEL)+2))]
[88] PLABEL+((pXLABEL),1)pXLABEL
[89] OL OUT ',PLABEL+,PLABEL+(10 4 EFT PLABEL),((pXLABEL),6)p',
[90] +0

```

9


```

MULTIPLY[ ]
[1]  D←D\1+C,6I3,C+6I 3 120
[2]  MSGS 'OFF'
[3]  +(R90)/PL2,ST+6PK+R+0
[4]  (SM[2]-P,K+H[1]) TICMARKOP[2]+,
[5]  PL2:L+(1,QLA,MM+0=(SM[2]+1 2)O,|,H[2])\2
[6]  8 TICMARKOPL3+PL3+1+HS,L2+P×1Q+~HS,PCH[1]
[7]  L5+P-1+HS,L4+Q×1~P2,P×M+(P)1]-1-2 1
[8]  L1+((HS,HS×A+Q×Q8+P-8)A~A)/PL4+ 2 1
[9]  TM+TM,[1.5] TM+ 1+TM
[10] PL3:E+~I+0A×L+L×R
[11] +(L1×1Q<PL[D+1+D/X[:2]]+J+2+(D+X[:1]=N+|K-C)/A),L2
[12] +E,L[T]+(E\L)[T+(RALS1)/,PL]
[13] D+(E/,PE)[T,(U+L[T]>2)/T+D]
[14] L[(~U)/T]+(~U,U/1)/J+J,U/M+1
[15] PL4:+(A×12PD)/E,E+1V×L+L
[16] +(L5×1~V/U+(P2V×J=1QJ+1,J[T])^D=1QD+0,D[T+U[AD[U+AJ]]],L4
[17] +(A/U+1=+/(D×,=D+U/D)^J×,=J)/Q,J[T/,P×J+U/J]+(T+U/~T)/M-1
[18] U+1+(T=1QJ+1,J+J[T])×U=1QD+0,D+D[T+U[AD[U+AJ]]]
[19] Q:~P×1Q8<PT+~(1+T+P)I)∈D+(D-1),D+D+2×1T+P×D+U/D
[20] I←T\I
[21] I[D]+(U/M),U/J
[22] L+I[(E+~×I)\L
[23] E:~(×N)/PL5+V/T+(2 1 ×N)×,TM
[24] +(PL5+1),L+L[E\1+1,V×HM
[25] PL5:L+L[0,T,(( 4+P×L)PGLA1+T),0
[26] PT[TM[:1],N:],P[1+L]
[27] +(Q×R+×C+G-1)/L3
[28] (SM[2]-1) TICMARK-R90
[29] →U+(ST[3 4],1)/ 1 3 4 +I26
[30] 'SCALE FACTOR FOR ORDINATE: ',10×ST[5]
[31] →U+1+U
[32] 'SCALE FACTOR FOR ABSCISSA: ',10×ST[6]
[33] MSGS 'ON'

```



```

V DFT[ ] V
Z+W DFT X;D;E;F;G;H;I;J;K;L;Y
D+ 0123456789.-
[1] + (V/H≠[W+,W+(H+0)×L+1<ρX)/(DFTERR+0×F+2
[2] →(3 2 1 <ρX)/(DFTERR+F+0), 2 3 +I26
[3] + (2+I26),ρX+((V/ 1 2 =ρW)φ 1 2)Q(1,ρ,X)ρX
[4] X+(0 1 /ρX)ρX
[5] →((∧/(ρW)× 1 2, 2×E+1ρφX), 1×ρW)/(DFTERR×F+1), 3+I26
[6] I+1+I/0,,{10⊙{X+1>|X
[7] W+(2+I+W+(W≠0)+V/,X<0),W
[8] →(V/2>-/[1] W+Q(E,2)ρW)/(DFTERR+0×F+2
[9] Z+((K+1ρX),+/W[1;])ρ
[10] X+10.5+X×10*(ρX)ρW[2;]
[11] DFTLP:→(E<H+H+1)/DFTEND
[12] J+1+10|(|Y+X[;H])°.×10*-1+φ1I+W[1;H]
[13] J+((,J)×G+,Q(φJ)ρ(,Q(J≠1)V.∧(1I)°.≤1I-F+1),(K×1+F+W[2;H])ρ1
[14] + (∧/0≤Y)/2+I26
[15] J[1+(ρJ)]-1+(I-+/(K,I)ρG)+I×-1+1K]+12×Y<0
[16] J+(K,I)ρJ
[17] →(0=F)/3+I26
[18] J+J[;(1φ1G),(G+-/W[;H])+1F]
[19] J[;G]+11
[20] →DFTLP,ρZ[;(+/W[1;H-1])+1I]+D[1+J]
[21] DFTEND:→L/0
[22] →0×ρZ+,Z
[23] DFTERR: DFT , (3 6 ρ, RANK LENGTHDOMAIN')[F+1;], ' PROBLEM. '
V
VAPLNAME[ ] V
FID+APLNAME A;K;REM
A REMOVE EXTRA BLANKS
A+1+(K\1φK+ , '≠A)/A+ , 'A, '
A FIND END OF FILENAME
K+(A, ' )-11
A IF ONE WORD - SYNTAX ERROR
→(K=ρA)/ER1
A EXTRACT FILENAME
FID+8+K+A
A AND REMAINDER
REM+(K+1)+A
A FIND END OF FILETYPE
K+(REM, ' )-11
A ADD FILETYPE TO FILE
FID+FID,(8+K+REM)
A EXTRACT 2ND REMAINDER
REM+(K+1)+REM
A CHECK SPECIAL MODES
→((∧/'SY'=2+REM)V(∧/'* '=2+REM))/L1
A MODELETTER='P', UNLESS OTHERWISE
FID+FID,'ABCTP','ABCT',1+REM]
A MODENO='1', UNLESS OTHERWISE
FID+FID,'0234561','023456',1+1+REM]
→L2
L1:FID+FID,2+REM
A RECTYPE='F', UNLESS V SPECIFIED
L2:FID+FID, , 'FV'[(V'=1+REM)+1]
A CONVERT TO EBCDIC INTEGER
FID+2 OF FID
→0
[30] ER1: FILETYPE MISSING'
V
V AUTOSCALE[ ] V
C+C+(X[1;]+X[1;]=0)×0=C+((1/X)-D+1/X
F+F;G+10×[10×F+|C+H+SM×|((6,PI)×2ρA)+SM+16(SMf 1 5
F+G×ST[+/F°.≤ST+10,ST[V2T]]
X+(Cρ 0 0.5)+(Cρ;F)×X-(C+ρX)ρG+G×(0<G-C+C)|D+C+F×SM+2
V

```



```

VMPLOT[[]]V
V A MPLOT X;C;D;F;G;H;P;P2;HS;A;ST
[1] INITIAL
[2] AUTOSCALE
[3] SETAP
[4] MULTIPLY H+EM*(1+X+X)*EM
V
[1] OL OUT R;I;J;MAX
[2] +(2=ρR)/L1
[3] R+(1,ρR)ρR
[4] L1:→(OL=1)/OFF
[5] +0,ρ[+R
[6] OFF:MAX+1+ρR+*,R
[7] J+20f-1+ρR
[8] I+1
[9] L2:(J+R[I;]) WRITE APLN
    →(MAX≥I+1)/L2
V
[1] INITIAL
[2] →(0=×/(2ρA),D+ρX), 2 1 <ρX)/0,PL2- 1 0
[3] +PL2,D+ρX+(1ρ,X),[1.5] X
[4] X+(D+2+D)ρX
[5] PL2:X+R90φ(.Q 0 1 +X),[1.5](C+×/D+D- 0 1)ρX[;1]
V
VAND[[]]V
L+A AND B;C;D
→((2<ρA)∨3<ρB),0×ρB)/ 17 3
[1] B+,B
[2] →((3=ρB)∧1×1ρB),2=ρA)/ 17 7
[3] A+,A
[4] +/(1ρA)×1,D),1×D+1ρ-2φB)/16
[5] A+((D×ρA)∣D(ρA),1)ρA
[6] +(1×ρB)/9
[7] B+((1ρB){(1=ρB)×1ρA),1)ρB
[8] +((1/D×1,1ρA),1×D+1ρ-2φB)/ 16 11
[9] B+((3=ρB)ρ1),(1ρA),1ρφB)ρB
[10] +(3=ρB)/14
[11] L+((C+1ρφA)ρ0),(1ρφB)ρ1)\B
[12] +0×ρL[;1C]+A
[13] L+(1,((C+1ρφA)ρ0),(-1+1ρφB)ρ1)\B
[14] +0×ρL[;1+1C]+A
[15] →0=ρ[+ ARGUMENTS OF AND ARE NOT CONFORMABLE.
[16] →0=ρ[+ ARGUMENTS OF AND IS OF IMPROPER RANK.
[17] *AN ARGUMENT OF AND IS OF IMPROPER RANK.
V
VTOT[[]]V
TOT
[1] ARRAY[1;]+,BRRAY[1+,(A-1);]
[2] K+1
[3] TAB1:K+K+1
[4] ARRAY[K;]+((,BRRAY[K+(A-K);]),(,BRRAY[1(K-1);]))
[5] →(K=(A-1))/TAB2
[6] →TAB1
[7] TAB2:ARRAY[A;]+,BRRAY[1(A-1);]
[8] →0
V

```



```

VTICMARK[ ]V
V U TICMARK ISV;C;I;J;L;T;NB;VT;O;N;E;K
[1] +(PL3-1~K+R90<ISV),BPx1~ISV< 0 1
[2] (xρ,TT)/EΣ[2],((Of[0.5x8+H[2]]-ρ,TT)ρ', ),TT,1+EΣ
[3] ((R90<O[8]+[0.5xH[2]]-ρ,TT)ρ', ),TT
[4] C+ρNB+G[J]+F[J]xTM+0,ΣM[J]x1[H[J]]ΣM[J+2-ISO]
[5] PL3;N+(C,VT+v/N)ρ(N+O>PT+[NBx10*U-I+[/O+1[1+E+{10●|NB+NB=0}\'-,
[6] L+(U+U-N+-VT)+1-VT+(φ'0'Λ,=PT+N, '0123456789,[1+Q(Uρ10)T|PT])10
[7] +(U>T+VT+f/I,(L+L≠I),(ISO)x2+L-I),L≥U-VT+L>I)/ 3 2 +I26
[8] +(I26)-I+ρ,ST[6-K]+I-1
[9] →PL3,ρNB+(10*-L)x[0.5+NBx10*(L+3+U+U-VT)-ST[6-K]+[//(NB≠0)/E
[10] PT+((~(U+J)ε(I+J+VT),1-1+J+U-T)\(C,U)ρ,(VTφO)\(.O+(I-O)°.<NφIU)/.PT
[11] ST[4-K]+ST[6-K]≠ρPT[;I+J+VT]+',
[12] +ISV/OxρPT+((C+1),~VT)ρ', ),(1-1+C,U)1PT
[13] (ΣM[2]-9)φ', ',PT
[14] →R90/O
[15] BP:(((~R90)xO[8+10.5xH[2]]-ρ,BL)ρ', ),BL
[16] (xρ,BT)/EΣ[2],((Of[0.5x8+H[2]]-ρ,BT)ρ', ),BT,1+EΣ
V

```

```

VSETΔAP[ ]V
V SETΔAP
[1] D+ρA+(A>A+0)/A+A,C-+/A+2+A,D[2]ρD[1]
[2] +(D>1)/4,ρP+',_1',(DρEC),((P2+(xρ,EC2)Λ~HS+HΣΛ~R90)+EC2),1+EΣ
[3] →A+ρA+~P2
[4] A+1+/(1C)°.>(D°,>D+1D)+.xA
V

```

```

VVS[ ]V
V M+A VS B;C;D
[1] +(((ρρB+,B)<ρρB), 2 1 0 <ρρA)/ 8 8 4 3
[2] A+((ρB),1)ρA
[3] A+((x/ρA),1)ρA
[4] +(Λ/(ρB)≠1,1ρρA)/9
[5] M+(0,(1ρφρA)ρ1)\Λ
[6] M[;11]+B
[7] →0xρρM+(1,ρM)ρM
[8] →0=ρ[+,'AN ARGUMENT OF VS IS OF IMPROPER RANK.',
[9] 'ARGUMENTS OF VS ARE NOT CONFORMABLE.',
V

```


LIST OF REFERENCES

1. Chatfield, C. and P. Sprent. "Convenient Probability Plotting Positions for the Normal Distribution," Applied Statistics 25: 47-50, 1976.
2. Dixon, M. W. A Statistical Analysis of Deviation from Target Cost in NAVAIRSCOMHQ Fixed-Price Incentive Contracts during the 1949-1965 Time Frame. U.S. Naval Postgraduate School, Monterey, California, 1973.
3. Fishman, G. S. Concepts and Methods of Discrete Event Digital Simulation. John Wiley and Sons, Inc., New York, 1973.
4. Gillman, L. and A. J. Rose. APL An Interactive Approach. John Wiley and Sons, Inc., New York, 1974.
5. Lewis, P.A.W. "A Branching Poisson Process Model for the Analysis of Computer Failure Patterns." The Journal of the Royal Statistical Society Series B (Methodological). 26: 398-456, 1964.
6. Lewis, P.A.W. and D. R. Cox. "A Statistical Analysis of Telephone Circuit Error Data." IEEE Transactions on Communications Systems, Com-14, 382-389, 1966.
7. Lewis, P.A.W. and G. S. Shedler. "Empirically Derived Micromodels for Sequences of Page Exception." IBM Journal of Research and Development 17: 89-90, 1973.
8. Lewis, P.A.W., L. H. Liu, D. W. Robinson and M. Rosenblatt. Empirical Sampling Study of a Goodness of Fit Statistic for Density Function Estimation. U. S. Naval Postgraduate School, Monterey, California, 1975.
9. Miller, R. G. "The Jackknife - A Review." Biometrika 61: 1, 1974.
10. Pearson, E. S. and H. W. Hartley. Biometrika Tables for Statisticians, v. I. Cambridge University Press, Cambridge, 1966.
11. Rosenblatt, M. "Curve Estimates." The Annals of Mathematical Statistics, v. 27, 1956.
12. Weast, R. C. C.R.C. Standard Mathematical Tables, 19th Ed., The Chemical Rubber Co., Ohio, 1971.
13. Wilk, A. and B. Gnanadesikan. "Probability Plotting Methods for the Analysis of Data." Biometrika 55: 1-17, 1968.

INITIAL DISTRIBUTION LIST

	No. Copies
1. Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2. Chief, Civil Schools Branch Department of Army U. S. Army Military Personnel Center Alexandria, Virginia 22332	1
3. Library, Code 0142 Naval Postgraduate School Monterey, California 93940	2
4. Dr. P.A.W. Lewis, Code 55Lw Department of Operations Research Naval Postgraduate School Monterey, California 93940	15
5. Lt. D. R. Bouchoux, Code 55Bh Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
6. Dr. R. W. Butterworth, Code 55Bd Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
7. Dr. D. T. Gavor, Code 55Gv Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
8. Dr. K. T. Marshall, Code 55Mt Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
9. Dr. R. R. Read, Code 55Re Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
10. Capt. Dennis R. Hutchinson 1333 Capistrano Drive Salt Lake City, Utah 84116	2

3 OCT 78
31 DEC 80
18 JUN 81

25021
26405
26405

Thesis
H964 Hutchinson
c.1 APL histogram, density
estimation and probability
plotting routines.

167935

3 OCT 78
31 DEC 80
18 JUN 81

25021
26405
26405

Thesis
H964 Hutchinson
c.1 APL histogram, density
estimation and probability
plotting routines.

167935

thesH964

APL histogram, density estimation and p



3 2768 001 03600 7

DUDLEY KNOX LIBRARY